**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, C12N 15/11, C07K 16/28, G01N 33/68

A2

(11) International Publication Number: WO 96/37610

(43) International Publication Date: 28 November 1996 (28.11.96)

(21) International Application Number: PCT/IB96/00696

(22) International Filing Date: 24 May 1996 (24.05.96)

(30) Priority Data:
08/452,630 25 May 1995 (25.05.95) US

(71) Applicant: MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V. [DE/DE]; Hofgartenstrasse 2, D-80539 München (DE).

(72) Inventors: ULLRICH, Axel; 70 Palmer Lane, Portola Valley, CA 94028 (US). MOSSIE, Kevin; 980 Kiely Boulevard #311, Santa Clara, CA 95051 (US). ALVES, Frauke; Rohnsweg 2, D-37085 Göttingen (DE).

(74) Agent: SCHLEE, Alexander; Viering, Jentschura & Partner, Steinsdorfstrasse 6, D-80538 München (DE).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG). European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**
*Without international search report and to be republished upon receipt of that report.*

(54) Title: CCK-4, A RECEPTOR TYROSINE KINASE, AND METHODS FOR DIAGNOSIS AND TREATMENT OF CCK-4 SIGNAL TRANSDUCTION DISORDERS

(57) **Abstract**

The present invention relates to CCK-4 polypeptides, nucleic acids encoding such polypeptides, cells, tissues and animals containing such nucleic acids, antibodies to such polypeptides, assays utilizing such polypeptides, and methods relating to all of the foregoing. Methods for treatment, diagnosis, and screening are provided for diseases or conditions characterized by an abnormality in a signal transduction disorder. The signal transduction pathway involves an interaction between a CCK-4 receptor tyrosine kinase and a receptor for the kinase.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AM** | Armenia | **GB** | United Kingdom | **MW** | Malawi |
| **AT** | Austria | **GE** | Georgia | **MX** | Mexico |
| **AU** | Australia | **GN** | Guinea | **NE** | Niger |
| **BB** | Barbados | **GR** | Greece | **NL** | Netherlands |
| **BE** | Belgium | **HU** | Hungary | **NO** | Norway |
| **BF** | Burkina Faso | **IE** | Ireland | **NZ** | New Zealand |
| **BG** | Bulgaria | **IT** | Italy | **PL** | Poland |
| **BJ** | Benin | **JP** | Japan | **PT** | Portugal |
| **BR** | Brazil | **KE** | Kenya | **RO** | Romania |
| **BY** | Belarus | **KG** | Kyrgystan | **RU** | Russian Federation |
| **CA** | Canada | **KP** | Democratic People's Republic of Korea | **SD** | Sudan |
| **CF** | Central African Republic | | | **SE** | Sweden |
| **CG** | Congo | **KR** | Republic of Korea | **SG** | Singapore |
| **CH** | Switzerland | **KZ** | Kazakhstan | **SI** | Slovenia |
| **CI** | Côte d'Ivoire | **LI** | Liechtenstein | **SK** | Slovakia |
| **CM** | Cameroon | **LK** | Sri Lanka | **SN** | Senegal |
| **CN** | China | **LR** | Liberia | **SZ** | Swaziland |
| **CS** | Czechoslovakia | **LT** | Lithuania | **TD** | Chad |
| **CZ** | Czech Republic | **LU** | Luxembourg | **TG** | Togo |
| **DE** | Germany | **LV** | Latvia | **TJ** | Tajikistan |
| **DK** | Denmark | **MC** | Monaco | **TT** | Trinidad and Tobago |
| **EE** | Estonia | **MD** | Republic of Moldova | **UA** | Ukraine |
| **ES** | Spain | **MG** | Madagascar | **UG** | Uganda |
| **FI** | Finland | **ML** | Mali | **US** | United States of America |
| **FR** | France | **MN** | Mongolia | **UZ** | Uzbekistan |
| **GA** | Gabon | **MR** | Mauritania | **VN** | Viet Nam |

# DESCRIPTION

## CCK-4, A RECEPTOR TYROSINE KINASE AND METHODS FOR DIAGNOSIS AND TREATMENT OF CCK-4 SIGNAL TRANSDUCTION DISORDERS

### Field of the Invention

The present invention relates to the novel protein CCK-4 nucleotide sequences encoding CCK-4 as well as various products and methods useful for the diagnosis and treatment of various CCK-4 related diseases and conditions associated with abnormal cellular signal transduction pathways.

### Background of the Invention

None of the following is admitted to be prior art to the invention.

Cellular signal transduction is a fundamental mechanism whereby external stimuli that regulate diverse cellular processes are relayed to the interior of cells. One of the key biochemical mechanisms of signal transduction involves the reversible phosphorylation of tyrosine residues on proteins. The phosphorylation state of a protein is modified through the reciprocal actions of tyrosine kinases (TKs) and tyrosine phosphatases (TPs).

Receptor tyrosine kinases (RTKs) belong to a family of transmembrane proteins and have been implicated in cellular signaling pathways. The predominant biological activity of some RTKs is the stimulation of cell growth and proliferation, while other RTKs are involved in arresting growth and promoting differentia-

tion. In some instances, a single tyrosine kinase can inhibit, or stimulate, cell proliferation depending on the cellular environment in which it is expressed.

RTKs are composed of at least three domains: an extracellular ligand binding domain, a transmembrane domain and a cytoplasmic catalytic domain that can phosphorylate tyrosine residues. Ligand binding to membrane-bound receptors induces the formation of receptor dimers and allosteric changes that activate the intracellular kinase domains and result in the self-phosphorylation (autophosphorylation and/or transphosphorylation) of the receptor on tyrosine residues. RTKs are also known to form heterodimers. A possible role for receptor heterodimerization is described in Carraway and Cantley, Cell 78:5-8 (1994). Their intrinsic tyrosine kinase is activated upon ligand binding, thereby initiating a complex signal transduction pathway that begins with receptor autophosphorylation and culminates in the tyrosine phosphorylation of a variety of cellular substrates and ultimately in the initiation of nuclear events necessary for the overall cell response. Individual phosphotyrosine residues of the cytoplasmic domains of receptors may serve as specific binding sites that interact with a host of cytoplasmic signaling molecules, thereby activating various signal transduction pathways.

In the case of heterodimers it is not necessary that both members of the heterodimer be enzymatically active. For example co-expression of the RTKs of Her2 and Her3 (which is enzymatically inactive)

produces a high-affinity receptor for herequlin which induces phosphorylation of the heterodimer. Sliwkowski et al., Jrnl Bio. Chem. 269:14661-14665 (1994). As Her3 has an abundance of potential tyrosine phosphorylation sites, it has been proposed that phosphorylated Her3 as a docketing protein for a number of downstream signaling molecules (see below). Constitutive activation/phosphorylation has been demonstrated in breast tumor cells. Krauss et al., PNAS 90:2900-2904 (1993).

The intracellular, cytoplasmic, non-receptor protein tyrosine kinases do not contain a hydrophobic transmembrane domain or an extracellular domain and share non-catalytic domains in addition to sharing their catalytic kinase domains. Such non-catalytic domains include the SH2 domains (SRC homology domain 2) and SH3 domains (SRC homology domain 3). The non-catalytic domains are thought to be important in the regulation of protein-protein interactions during signal transduction.

A central feature of signal transduction (for reviews, see Posada and Cooper, *Mol. Biol. Cell* 3:583-392, 1992; Hardie, *Symp. Soc. Exp. Biol.* 44:241-255, 1990), is the reversible phosphorylation of certain proteins. Receptor phosphorylation stimulates a physical association of the activated receptor with target molecules. Some of the target molecules such as phospholipase Cγ are in turn phosphorylated and activated. Such phosphorylation transmits a signal to the cytoplasm. Other target molecules are not phosphorylated, but assist in signal transmission by

acting as adapter molecules for secondary signal transducer proteins. For example, receptor phosphorylation and the subsequent allosteric changes in the receptor recruit the *Grb*-2/SOS complex to the catalytic domain of the receptor where its proximity to the membrane allows it to activate *ras*.

The secondary signal transducer molecules generated by activated receptors result in a signal cascade that regulates cell functions such as cell division or differentiation. Reviews describing intracellular signal transduction include Aaronson, *Science*, *254*:1146-1153, 1991; Schlessinger, *Trends Biochem. Sci.*, *13*:443-447, 1988; and Ullrich and Schlessinger, *Cell*, 61:203-212, 1990.

Since several RTKs and growth factors were originally identified as activated oncogenes (Aaronson, 1991; Bishop, 1991), there has always been a belief that some RTKs may be involved in the development of some cancers (Bishop, 1991). Several studies now appear to support this notion (Aaronson, 1991; see Plowman et al., 1994). These include the high correlation of RTK overexpression with certain human cancers, including HER2 with breast and ovarian cancers (Slamon et al., 1987), PDGF and its receptors with a high fraction of sarcomas and glially derived neoplasms, and EGF-R with squamous cell carcinomas and glioblastomas (reviewed in Aaronson et al., 1991; Plowman et al., 1994).

Colorectal cancer has served as a model for deciphering the molecular alterations involved in the development of human neoplasms. This process has been

characterized by the activation of cellular oncogenes, coupled with the inactivation of tumor suppressor genes. Positional mapping efforts have led to the identification of several tumor suppressor genes that are frequently deleted or mutated during the neoplastic progression of this cancer. The candidate colorectal tumor suppressor genes include APC, p53, DCC, MCC and the DNA mismatch repair genes, hMSH2 and hMLIII (Bronner et al., Nature 368:258-261, 1994; Kinzler et al., Science 253:661-665, 1991; Groden et al., Cell 66:589-600, 1991; Fearon and Vogelstein, Cell 61:759-767, 1990).

Conversely, activation of cellular oncogenes is felt to play an additional role in the pathogenesis of colorectal cancer. The most frequent event is the activation of Ki-ras which occurs in 40-50% of these tumors (Bos, Nature 327:293-297, 1987; Pories et al., Oncogene, 7:885-893, 1992). Amplification or overexpression of receptor tyrosine kinases (RTKs) such as EGF receptor and HER2 have also been implicated in the progression of colorectal cancer (Mendelsohn, Monogr. Natl. Cancer Inst., 13:125-131, 1992; D'Emilia et al., Surg. Oncol., 1:97-105, 1989). Similarly, other RTKs including Met, FGF receptor 3, IGF-I receptor, c-Kit and Flk have been observed to be overexpressed in this tumor type (Faletto, EXS, 65:107-130, 1993; Murgue et al., Cancer Res., 54:5206-5211, 1994; Yamori et al., Gan. To Kagaku Ryoho, 20:393-398, 1993; Toyota et al., Tumour Biol., 14:295-302, 1993; Brown et al., Cancer Res., 53:4727-4735, 1993).

## Summary of the Invention

The present invention relates to CCK-4 polypeptides, nucleic acids encoding such polypeptides, cells containing such nucleic acids, antibodies to such polypeptides, assays utilizing such polypeptides, and methods relating to all of the foregoing. In particular, this invention relates to methods for diagnosis and treatment of a disorder, preferably a disorder characterized by an abnormality in a signal transduction pathway, wherein the signal transduction pathway involves the CCK-4 receptor tyrosine kinase. As used herein, the term CCK-4 receptor tyrosine kinase is meant to include receptor tyrosine kinase-like genes that lack detectable tyrosine kinase activity as is explained in detail below.

The present invention is based upon the isolation and characterization of a new RTK which we have designated (CCK-4). CCK-4 was found using a polymerase chain reaction (PCR) based approach and surprisingly seems to be preferentially expressed in cancerous colon tissue in comparison with normal colon. We have determined that modulation CCK-4 receptor tyrosine kinase cellular signaling is useful in therapeutic procedures. CCK-4 polypeptides are involved in various signal transduction pathways and thus the present invention provides several agents and methods useful for diagnosing, treating, and preventing various diseases or conditions associated with abnormalities in these pathways.

One example of the utility of the present invention is the ability to diagnosis colon cancer, and/or a particular subset of colon cancers by detecting the presence or amount of a CCK-4 polypeptide or nucleic acid encoding such a polypeptide.

Thus, in a first aspect the invention features an isolated, enriched, or purified nucleic acid encoding a CCK-4 polypeptide.

By "isolated" in reference to nucleic acid is meant a polymer of 2 (preferably 21, more preferably 39, most preferably 75) or more nucleotides conjugated to each other, including DNA or RNA that is isolated from a natural source or that is synthesized. The isolated nucleic acid of the present invention is unique in the sense that it is not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only nucleotide chain present, but that it is essentially free (about 90 - 95% pure at least) of non-nucleotide material naturally associated with it and thus is meant to distinguish from isolated chromosomes.

By the use of the term "enriched" in reference to nucleic acid is meant that the specific DNA or RNA sequence constitutes a significantly higher fraction (2 - 5 fold) of the total DNA or RNA present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This

could be caused by a person by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a combination of the two. However, it should be noted that enriched does not imply that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased. The term significant here is used to indicate that the level of increase is useful to the person making such an increase, and generally means an increase relative to other nucleic acids of about at least 2 fold, more preferably at least 5 to 10 fold or even more. The term also does not imply that there is no DNA or RNA from other sources. The other source DNA may, for example, comprise DNA from a yeast or bacterial genome, or a cloning vector such as pUC19. This term distinguishes from naturally occuring events, such as viral infection, or tumor type growths, in which the level of one mRNA may be naturally increased relative to other species of mRNA. That is, the term is meant to cover only those situations in which a person has intervened to elevate the proportion of the desired nucleic acid.

It is also advantageous for some purposes that a nucleotide sequence be in purified form. The term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation); instead, it represents an indication that the sequence is relatively purer than in the natural environment (compared to the natural level this level should be at

least 2-5 fold greater, *e.g.*, in terms of mg/ml). Individual clones isolated from a cDNA library may be purified to electrophoretic homogeneity. The claimed DNA molecules obtained from these clones could be obtained directly from total DNA or from total RNA. The cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process which includes the construction of a cDNA library from mRNA and isolation of distinct cDNA clones yields an approximately $10^6$-fold purification of the native message. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated.

By "a CCK-4 polypeptide" is meant 25 (preferably 30, more preferably 35, most preferably 40) or more contiguous amino acids set forth in the full length amino acid sequence of Figure 1, or a functional derivative thereof as described herein. The CCK-4 polypeptide can be encoded by a full-length nucleic acid sequence or any portion of the full-length nucleic acid sequence, so long as a functional activity of the polypeptide is retained. Such functional activity can be, for example, the formation of homo- or heterodimers,

and/or providing binding sites for intracellular signaling components.

In preferred embodiments the isolated nucleic acid comprises, consists essentially of, or consists of a nucleic acid sequence set forth in the full length nucleic acid sequence of Figure 1, a functional derivative thereof, or at least 75, 90, 105, 120 or 150 contiguous nucleotides thereof; the CCK-4 polypeptide comprises, consists essentially of, or consists of at least 25, 30, 35, or 40 contiguous amino acids of a CCK-4 polypeptide. The nucleic acid may be isolated from a natural source by cDNA cloning or subtractive hybridization; the natural source may be mammalian (human) blood, semen, or tissue and the nucleic acid may be synthesized by the triester method or by using an automated DNA synthesizer. In yet other preferred embodiments the nucleic acid is a conserved or unique region, for example those useful for the design of hybridization probes to facilitate identification and cloning of additional polypeptides, the design of PCR probes to facilitate cloning of additional polypeptides, and obtaining antibodies to polypeptide regions.

By "conserved nucleic acid regions", are meant regions present on two or more nucleic acids encoding a CCK-4 polypeptide, to which a particular nucleic acid sequence can hybridize under lower stringency conditions. Examples of lower stringency conditions suitable for screening for nucleic acid encoding CCK-4 polypeptides are provided in Abe, *et al.* J. Biol. Chem., 19:13361 (1992) (hereby incorporated by reference herein

in its entirety, including any drawings). Preferably, conserved regions differ by no more than 25 out of 20 nucleotides.

By "unique nucleic acid region" is meant a sequence present in a full length nucleic acid coding for a CCK-4 polypeptide that is not present in a sequence coding for any other naturally occurring polypeptide. Such regions preferably comprise 30 or 45 contiguous nucleotides present in the full length nucleic acid encoding a CCK-4 polypeptide. In particular, a unique nucleic acid region is not present in KLG and is preferably of mammalian origin.

The invention also features a nucleic acid probe for the detection of a CCK-4 polypeptide or nucleic acid encoding a CCK-4 polypeptide in a sample. The nucleic acid probe contains nucleic acid that will hybridize to a sequence set forth in Figure 1 or a functional derivative thereof.

In preferred embodiments the nucleic acid probe hybridizes to nucleic acid encoding at least 12, 75, 90, 105, 120 or 150 contiguous amino acids of the full-length sequence set forth in Figure 1 or a functional derivitive thereof. Various low or high stringency hybridization conditions may be used depending upon the specificity and selectivity desired. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides.

Methods for using the probes include detecting the presence or amount CCK-4 RNA in a sample by contacting the sample with a nucleic acid probe under conditions such that hybridization occurs and detecting the presence or amount of the probe bound to CCK-4 RNA. The nucleic acid duplex formed between the probe and a nucleic acid sequence coding for a CCK-4 polypeptide may be used in the identification of the sequence of the nucleic acid detected (for example see, Nelson *et al.*, in Nonisotopic DNA Probe Techniques, p. 275 Academic Press, San Diego (Kricka, ed., 1992) hereby incorporated by reference herein in its entirety, including any drawings). Kits for performing such methods may be constructed to include a container means having disposed therein a nucleic acid probe.

The invention also features recombinant nucleic acid, preferably in a cell or an organism. The recombinant nucleic acid may contain a sequence set forth in Figure 1 or a functional derivative thereof and a vector or a promoter effective to initiate transcription in a host cell. The recombinant nucleic acid can alternatively contain a transcriptional initiation region functional in a cell, a sequence complimentary to an RNA sequence encoding a CCK-4 polypeptide and a transcriptional termination region functional in a cell.

In another aspect the invention features an isolated, enriched, or purified CCK-4 polypeptide.

By "isolated" in reference to a polypeptide is meant a polymer of 2 (preferably 7, more preferably 13,

most prefereably 25) or more amino acids conjugated to each other, including polypeptides that are isolated from a natural source or that are synthesized. The isolated polypeptides of the present invention are unique in the sense that they are not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only amino acid chain present, but that it is essentially free (about 90 - 95% pure at least) of non-amino acid material naturally associated with it.

By the use of the term "enriched" in reference to a polypeptide is meant that the specific amino acid sequence constitutes a significantly higher fraction (2 - 5 fold) of the total of amino acids present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other amino acids present, or by a preferential increase in the amount of the specific amino acid sequence of interest, or by a combination of the two. However, it should be noted that enriched does not imply that there are no other amino acid sequences present, just that the relative amount of the sequence of interest has been significantly increased. The term significant here is used to indicate that the level of increase is useful to the person making such an

increase, and generally means an increase relative to other amino acids of about at least 2 fold, more preferably at least 5 to 10 fold or even more. The term also does not imply that there is no amino acid from other sources. The other source amino acid may, for example, comprise amino acid encoded by a yeast or bacterial genome, or a cloning vector such as pUC19. The term is meant to cover only those situations in which man has intervened to elevate the proportion of the desired nucleic acid.

It is also advantageous for some purposes that an amino acid sequence be in purified form. The term "purified" in reference to a polypeptide does not require absolute purity (such as a homogeneous preparation); instead, it represents an indication that the sequence is relatively purer than in the natural environment (compared to the natural level this level should be at least 2-5 fold greater, *e.g.*, in terms of mg/ml). Purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. The substance is preferably free of contamination at a functionally significant level, for example 90%, 95%, or 99% pure.

In preferred embodiments the CCK-4 polypeptide contains at least 25, 30, 35, 40, or 50 contiguous amino acids of the full-length sequence set forth in Figure 1, or a functional derivitive thereof.

In yet another aspect the invention features an antibody (e.g., a monoclonal or polyclonal antibody)

having specific binding affinity to a CCK-4 polypeptide. The antibody contains a sequence of amino acids that is able to specifically bind to a CCK-4 polypeptide. By "specific binding affinity" is meant that the antibody binds to CCK-4 polypeptides with greater affinity than it binds to other polypeptides under specified conditions.

Antibodies having specific binding affinity to a CCK-4 polypeptide may be used in methods for detecting the presence and/or amount of a CCK-4 polypeptide in a sample by contacting the sample with the antibody under conditions such that an immunocomplex forms and detecting the presence and/or amount of the antibody conjugated to the CCK-4 polypeptide. Diagnostic kits for performing such methods may be constructed to include a first container means containing the antibody and a second container means having a conjugate of a binding partner of the antibody and a label.

In another aspect the invention features a hybridoma which produces an antibody having specific binding affinity to a CCK-4 polypeptide. By "hybridoma" is meant an immortalized cell line which is capable of secreting an antibody, for example a CCK-4 antibody. In preferred embodiments the CCK-4 antibody comprises a sequence of amino acids that is able to specifically bind a CCK-4 polypeptide.

Another aspect of the invention features a method of detecting the presence or amount of a compound capable of binding to a CCK-4 polypeptide. The method involves incubating the compound with a CCK-4

polypeptide and detecting the presence or amount of the compound bound to the CCK-4 polypeptide.

Thus, in another aspect, the invention features a method for treating a patient having a disease or condition characterized by an abnormality in a signal transduction pathway, wherein the signal transduction pathway involves the interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner. The disorder may also be characterized by an abnormal level of interaction between CCK-4 receptor tyrosine kinase and a CCK-4 binding partner. The method includes disrupting or promoting that interaction (or signal) in vivo. The method also involves inhibiting or promoting the activity of the complex formed between CCK-4 receptor tyrosine kinase and a CCK-4 binding partner.

By "disease or condition" is meant a state which is recognized as abnormal by members of the medical community. The disease or condition may be characterized by an abnormality in one or more signal transduction pathways in a cell, preferably a colon, blood, or cancer cell, wherein one of the components of the signal transduction pathway is a CCK-4 receptor tyrosine kinase.

Examples of diseases or conditions to be treated or diagnosed by the present invention include neuroproliferative disorders, cancers, (especially colon cancers, hyperproliferative disorders such as psoriasis and neurofibromatosis. These and other diseases or conditions are often characterized by one or more of the

following symptoms: tumors, astasia, aphasia, paralysis, paresea, and paralagies.

By "abnormality" is meant a level which is statistically different from the level observed in organisms not suffering from such a disease or condition and may be characterized as either an excess amount, intensity or duration of signal or a deficient amount, intensity or duration of signal. The abnormality in signal transduction may be realized as an abnormality in cell function, viability or differentiation state. We have determined that such abnormal interaction in a pathway can be alleviated by action at the CCK-4-binding partner interaction site in the pathway.

An abnormal interaction level may also either be greater or less than the normal level and may impair the normal performance or function of the organism. Thus, it is also possible to screen for agents that will be useful for treating a disease or condition, characterized by an abnormality in the signal transduction pathway, by testing compounds for their ability to affect the interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner, since the complex formed by such interaction is part of the signal transduction pathway. However, the disease or condition may be characterized by an abnormality in the signal transduction pathway even if the level of interaction between the CCK-4 receptor tyrosine kinase and a CCK-4 binding partner is normal.

By "interact" is meant any physical association between proteins, whether covalent or non-covalent.

Examples of non-covalent bonds include electrostatic bonds, hydrogen bonds, and Van der Waals bonds. Stryer, Biochemistry, 1988, pages 7-8. Furthermore, the interactions between proteins may either be direct or indirect. Another example of an indirect interaction is the independent production, stimulation, or inhibition of both CCK-4 receptor tyrosine kinase and a CCK-4 binding partner by a regulatory agent. Depending upon the type of interaction present, various methods may be used to measure the level of interaction. For example, the strengths of covalent bonds are often measured in terms of the energy required to break a certain number of bonds (i.e., kcal/mol) Non-covalent interactions are often described as above, and also in terms of the distance between the interacting molecules. Indirect interactions may be described in a number of ways, including the number of intermediary agents involved, or the degree of control exercised over the CCK-4 receptor tyrosine kinase relative to the control exercised over the CCK-4 binding partner.

By "CCK-4 receptor tyrosine kinase" is meant an amino acid sequence substantially similar to the sequence shown in Figure 1, or fragments thereof. A sequence that is substantially similar will have at least 90% identity (preferably at least 95% and most preferably 99-100%) to the sequence of Figure 1. In preferred embodiments the CCK-4 receptor tyrosine kinase contains an approximately 24 amino acid long terminal signal peptide, and an approximately 23 amino acid long hydrophobic sequence predicted to represent the RTK

transmembrane domain (approximately positions 703-726) located between the putative ligand binding extracellular domain (approximately 702 residues) and the cytoplasmic tyrosine kinase domains (approximately 344 residues). In addition, the putative extracellular domain may contain ten potential N-glycosylation sites (Asn-X-Ser/Thr) and cysteine residues which could form seven immunoglobulin-like loops, similar to what is found in the KLG RTK. The ATP binding site, commonly defined by Gx GxxG consensus GxSxxG and there may be a conservative K for R substitution in the HRDL motif within the tyrosine kinase domain. In addition the DFG consensus sequence may be changed to ALG (Fig 1B). These consensus sequence alterations appear to impair or abolish the autophosphorylation capacity of CCK-4.

By "identity" is meant a property of sequences that measures their similarity or relationship. Identity is measured by dividing the number of identical residues by the total number of residues and multiplying the product by 100. Thus, two copies of exactly the same sequence have 100% identity, but sequences that are less highly conserved and have deletions, additions, or replacements may have a lower degree of identity. Those skilled in the art will recognize that several computer programs are available for determining sequence identity.

By "CCK-4 binding partner" is meant an amino acid sequence that interacts with or binds a CCK-4 RTK. The term includes ligands and/or substrates for the CCK-4 kinase.

By "disrupt" is meant that the interaction between the CCK-4 receptor tyrosine kinase and a CCK-4 binding partner is reduced either by preventing expression of the CCK-4 receptor tyrosine kinase, or by preventing expression of the CCK-4 binding partner, or by specifically preventing interaction of the naturally synthesized proteins having these domains or by interfering with the interaction of the proteins.

By "promote" is meant that the interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner is increased either by increasing expression of a CCK-4 receptor tyrosine kinase, or by increasing expression of a CCK-4 binding partner, or by decreasing the dephosphorylating activity of the corresponding regulatory TP (or other phosphatase acting on other phosphorylated signalling components) by promoting interaction of the CCK-4 receptor tyrosine kinase and a CCK-4 binding partner or by prolonging the duration of the interaction. Many bivalent or polyvalent linking agents are useful in coupling polypeptides, such as an antibody, to other molecules. For example, representative coupling agents can include organic compounds such as thioesters, carbodiimides, succinimide esters, diisocyanates, glutaraldehydes, diazobenzenes and hexamethylene diamines. This listing is not intended to be exhaustive of the various classes of coupling agents known in the art but, rather, is exemplary of the more common coupling agents. (See Killen and Lindstrom 1984, J. Immunol. 133:1335-2549;

Jansen, F.K., et al. 1982, Immunological Rev. 62:185-216; and Vitetta et al., supra).

By "signal transduction pathway" is meant the sequence of events that involves the transmission of a message from an extracellular protein to the cytoplasm through a cell membrane. The signal ultimately will cause the cell to perform a particular function, for example, to uncontrollably proliferate and therefore cause cancer. Various mechanisms for the signal transduction pathway (Fry et al., *Protein Science*, 2:1785-1797, 1993) provide possible methods for measuring the amount or intensity of a given signal. Depending upon the particular disease associated with the abnormality in a signal transduction pathway, various symptoms may be detected. Those skilled in the art recognize those symptoms that are associated with the various diseases described herein. Furthermore, since some adapter molecules recruit secondary signal transducer proteins towards the membrane, one measure of signal transduction is the concentration and localization of various proteins and complexes. In addition, conformational changes that are involved in the transmission of a signal may be observed using circular dichroism and fluorescence studies.

In a related aspect the invention features a method for screening for an agent useful for treatment of such a disease or condition by assaying potential agents for the ability to disrupt or promote that interaction. The screening may also involve assaying potential agents for the ability to remove or reduce the

effect of an abnormality in a signal transduction pathway, wherein the signal transduction pathway contains a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner.

By "screening" is meant investigating an organism for the presence or absence of a property. The process may include measuring or detecting various properties, including the level of signal transduction and the level of interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner.

Useful agents for treatment of such diseases can be identified by standard screening protocols in which measurement of such interaction is determined. For example, such an agent may be a peptide which either comprises, consists of, or consists essentially of a CCK-4 receptor tyrosine kinase or, alternatively, a fragment thereof.

By "comprising" it is meant including, but not limited to, whatever follows the word "comprising". Thus, use of the term "comprising" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present. By "consisting of" is meant including, and limited to, whatever follows the phrase "consisting of". Thus, the phrase "consisting of" indicates that the listed elements are required or mandatory, and that no other elements may be present. By "consisting essentially of" is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action

specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements.

In preferred embodiments the screening involves looking for agonists or antagonists of a protein of interest, for example CCK-4 or a CCK-4 binding partner. The term agonist refers to agents that bind the protein and that maintain the activity of the protein to which they bind. An antagonist competes with the natural ligand for binding the protein, but does not maintain the activity of the protein to which it binds.

Another aspect of the invention features a method for diagnosis of such a disease or condition. The method includes detecting the level of interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner.

By "diagnosis" is meant any method of identifying a symptom normally associated with a given disease or condition. Thus, an initial diagnosis may be conclusively established as correct by the use of additional confirmatory evidence such as the presence of other symptoms. Current classification of various diseases and conditions is constantly changing as more is learned about the mechanisms causing the diseases or conditions. Thus, the detection of an important symptom, such as the detection of an abnormal level of interaction between the CCK-4 receptor tyrosine kinases and binding partners

for the kinases may form the basis to define and diagnose a newly named disease or condition.

For example, conventional neurological diseases are classified according to the presence of a particular set of symptoms. However, a subset of these symptoms may both be associated with an abnormality in a particular signalling pathway, such as the *ras*$^{21}$ pathway and in the future these diseases may be reclassified as *ras*$^{21}$ pathway diseases regardless of the particular symptoms observed.

In other preferred embodiments the agent is therapeutically effective and has an $EC_{50}$ or $IC_{50}$ as described below. An $EC_{50}$ or $IC_{50}$ of less than or equal to 5 $\mu$M is preferable, and even more preferably less than or equal to 1 $\mu$M, 100 nmolar, 10 nmolar, or 1 nmolar. Such lower $EC_{50}$'s or $IC_{50}$'s are advantageous since they allow lower concentrations of molecules to be used *in vivo* or *in vitro* for therapy or diagnosis. The discovery of molecules with such low $EC_{50}$'s and $IC_{50}$'s enables the design and synthesis of additional molecules having similar potency and effectiveness. In addition, the molecule may have an $EC_{50}$ or $IC_{50}$ less than or equal to 5 $\mu$M at one or more, but not all cells chosen from the group consisting of colon cancer cell, parathyroid cell, bone osteoclast, juxtaglomerular kidney cell, proximal tubule kidney cell, distal tubule kidney cell, cell of the thick ascending limb of Henle's loop and/or collecting duct, central nervous system cell, keratinocyte in the epidermis, parafollicular cell in the thyroid (C-cell), intestinal cell, trophoblast in

the placenta, platelet, vascular smooth muscle cell, cardiac atrial cell, gastrin-secreting cell, glucagon-secreting cell, kidney mesangial cell, mammary cell, beta cell, fat/adipose cell, immune cell and GI tract cell.

By "therapeutically effective amount" is meant an amount of a pharmaceutical composition having a therapeutically relevant effect. A therapeutically relevant effect relieves to some extent one or more symptoms of the disease or condition in the patient; or returns to normal either partially or completely one or more physiological or biochemical parameters associated with or causative of the disease or condition. Generally, a therapeutically effective amount is between about 1 nmole and 1 $\mu$mole of the molecule, depending on its $EC_{50}$ or $IC_{50}$ and on the age and size of the patient, and the disease associated with the patient.

In another aspect the invention features a method for screening for human cells containing a CCK-4 RTK or an equivalent sequence. The method involves identifying the novel RTK in human cells using techniques that are routine and standard in the art, such as those described herein for identifying CCK-4 (e.g., cloning, Southern or Northern blot analysis, in situ hybridization, PCR amplification, etc.).

In preferred embodiments the method features screening cells involved in human cancers, such as colon cancer cells, for the presence of CCK-4. The invention also features methods of screening human cells for binding partners of CCK-4 RTKs and screening other

organisms for CCK-4 or the corresponding binding partner. In other preferred embodiments the agent is therapeutically effective and has an $EC_{50}$ or $IC_{50}$ as described herein.

In other aspects, the invention provides transgenic, nonhuman mammals containing a transgene encoding a CCK-4 polypeptide or a gene effecting the expression of a CCK-4 polypeptide. Such transgenic nonhuman mammals are particularly useful as an *in vivo* test system for studying the effects of introducing a CCK-4 polypeptide, regulating the expression of a CCK-4 polypeptide (i.e., through the introduction of additional genes, antisense nucleic acids, or ribozymes).

A "transgenic animal" is an animal having cells that contain DNA which has been artificially inserted into a cell, which DNA becomes part of the genome of the animal which develops from that cell. Preferred transgenic animals are primates, mice, rats, cows, pigs, horses, goats, sheep, dogs and cats. The transgenic DNA may encode for a human CCK-4 polypeptide. Native expression in an animal may be reduced by providing an amount of anti-sense RNA or DNA effective to reduce expression of the receptor.

In another aspect, the invention describes a polypeptide comprising a recombinant CCK-4 polypeptide or a unique fragment thereof. By "unique fragment," is meant an amino acid sequence present in a full-length CCK-4 polypeptide that is not present in any other naturally occurring polypeptide. Preferably, such a

sequence comprises 6 contiguous amino acids present in the full sequence. More preferably, such a sequence comprises 12 contiguous amino acids present in the full sequence. Even more preferably, such a sequence comprises 18 contiguous amino acids present in the full sequence.

By "recombinant CCK-4 polypeptide" is meant to include a polypeptide produced by recombinant DNA techniques such that it is distinct from a naturally occurring polypeptide either in its location (e.g., present in a different cell or tissue than found in nature), purity or structure. Generally, such a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

In another aspect, the invention describes a recombinant cell or tissue containing a purified nucleic acid coding for a CCK-4 polypeptide. In such cells, the nucleic acid may be under the control of its genomic regulatory elements, or may be under the control of exogenous regulatory elements including an exogenous promoter. By "exogenous" it is meant a promoter that is not normally coupled *in vivo* transcriptionally to the coding sequence for the CCK-4 polypeptide.

In another aspect, the invention features a CCK-4 polypeptide binding agent able to bind to a CCK-4 polypeptide. The binding agent is preferably a purified antibody which recognizes an epitope present on a CCK-4 polypeptide. Other binding agents include molecules which bind to the CCK-4 polypeptide and analogous molecules which bind to a CCK-4 polypeptide.

By "purified" in reference to an antibody is meant that the antibody is distinct from naturally occurring antibody, such as in a purified form. Preferably, the antibody is provided as a homogeneous preparation by standard techniques. Uses of antibodies to the cloned polypeptide include those to be used as therapeutics, or as diagnostic tools.

In another aspect the invention features a method for the diagnosis of colon cancer in a patient by detecting the presence or amount of a CCK-4 polypeptide or nucleic acid encoding a CCK-4 polypeptide. In a preferred embodiment the method of diagnosis is for a particular subset of colon cancers characterized by the presence or elevated amount of a CCK-4 polypeptide or nucleic acid encoding a CCK-4 polypeptide.

The summary of the invention described above is non-limiting and other features and advantages of the invention will be apparent from the following description of the preferred embodiments, and from the claims.

## Brief Decsription of the Drawings

Figure 1 illustates the consensus nucleotide sequence (4258 bp) and deduced amino acid sequence (1076 amino acids) of the full-length human CCK-4 gene. The consensus sequence was derived from eight cDNA clones isolated from three sources. The nucleotides are numbered at right and the amino acid sequence (one letter code) is shown from the putative ATG beginning at position 193 on the DNA map and the amino acid sequence position is shown below the nucleotide numbers. The

putative Kozak consensus sequence is underlined. The putative signal and transmembrane domain sequences are boldly underlined. The cysteine residues involved in the immunoglobulin domain loops are circled. These immunoglobulin domains conform to the consensus of a cysteine residue (circled) followed in 11 or 12 residues by a tryptophan and then another cysteine residue approximately 50 residues after the first cysteine located within the consensus DXGXYXC (SEQ. ID. NO. 1) (Chou and Hayman, 1991) and are located in residues 53-101, 150-200, 246-301, 343-391, 433-481, 524-570, 613-664. The putative N-glycosylation signals, NXS/T are underlined. Sequences resembling subdomains I, VI, and VII of tyrosine kinases are indicated by asterisks.

Figure 2 shows an amino acid comparison of the human CCK-4 gene with the chicken KLG gene. Homologous residues are boxed. The conserved N-glycosylation sites are indicated by . . . The position of the GxGxxG (subdomain I) and DFG (subdomain VII) conserved amino acid changes are indicated by asterisks.

## Description of the Preferred Embodiments

The present invention relates to CCK-4 polypeptides, nucleic acids encoding such polypeptides, cells, tissues and animals containing such nucleic acids, antibodies to such polypeptides, assays utilizing such polypeptides, and methods relating to all of the foregoing. RTKs were cloned using the established PCR-based approach to isolate RTK sequences that may be overexpressed in these colon carcinoma tumors.

Decribed here are the cloning and characterization of a novel human tyrosine-like gene, CCK-4, which belongs to the immunoglobulin superfamily of receptors and is a likely human homologue to the recently described chicken KLG gene (Chou and Hayman, 1991). CCK-4, like the chicken KLG gene, forms a unique class of receptor tyrosine kinase-like genes that lack detectable tyrosine kinase activity, suggesting a different mechanism of signal transduction for this class of molecules.

**I. Isolation and characterization of human CCK-4 cDNA**

In an effort to identify molecular defects related to the malignant characteristics of colon carcinoma, protein tyrosine kinase (PTK) expression in this tumor type was investgated by comprehensive polymerase chain reaction (PCR) analysis. PCR reactions were performed using a pool of poly $A^+$ RNA from human colon carcinoma cell lines and degenerate oligonucleotide primers encoding conserved amino acid sequences within the PTK catalytic domain (Hanks et al., 1988; Wilks, 1989; Alves et al., 1995) and mRNA preparations from primary colon carcinoma tumors. PCR fragments of the appropriate size were isolated, cloned, and characterized by nucleotide sequence analysis. Among 250 PCR clones analyzed, 24 different PTKs were identified, of which one represented a novel human receptor type tyrosine kinase (RTK), which was designated colon carcinoma kinase 4 (CCK-4).

The CCK-4 PCR fragment was used to screen a cDNA library prepared from human placenta mRNA. A positive clone containing a 2.2 kb cDNA insert was isolated first and subsequently used to isolate cDNAs encoding the entire CCK-4 RTK. The nucleotide sequence of CCK-4 cDNA analyzed from six (6) overlapping clones comprised 4,260 nucleotides (Figure 1). The first ATG codon is located at position 193, which, although no in-frame stop codons were detected upstream, is likely to represent the initiation codon since: (i) the nucleotide sequence upstream of this ATG is extremely G-C rich (ca. 76%), which is characteristic of 5' non-coding sequences of RTKs (Ullrich et al., 1984, 1985; Ebina et al., 1985; Shibuya et al. 1990), (ii) nucleotides flanking this ATG codon match Kozak's CC(A/G)CCATGG consensus for the translation initiation sites in mammals (Kozak, Nucl. Acids Res. 12:857-872, 1984), (REF), and (iii) is followed by 24 mostly hydrophobic residues, consistent with the characteristics of a signal peptide (REF) (Yarden et al., Annu. Rev. Biochem., 57:443-478, 1988). This ATG is part of an open reading frame of 3,210 nucleotides, which encode a 1,071 amino acid protein with the calculated molecular mass of 118,272, and are flanked by a 3' untranslated region of 475 nucleotides terminating with a putative polyadenylation signal (AATAA) at an appropriate distance from the beginning of the poly (A) tail.

With 70.2% amino acid sequence identity, CCK-4 is most closely related to the chicken KLG gene product,

a receptor tyorsine kinase with unknown function (Chou and Hayman, Proc. Natl. Acad. Sci. USA, 88:4897-4901, 1991).

II. ***CCK-4 gene expression in normal tissues and colon carcinoma***

Poly A$^+$ RNAs from either normal human tissue samples or colon carcinoma cell lines were analyzed by agarose gel electrophoresis under denaturing conditions and subsequent Northern blot hybridization analysis with $^{32}$P-labeled cDNA probes. While most normal human tissue RNAs contained no detectable CCK-4 gene transcripts, low levels of a hybridizing 4.4 kb mRNA were detected in thyroid gland and ovary with relatively high levels in the lung. No CCK-4 transcripts were detected in normal colon tissue mRNA even after extended exposure of Northern blots or extensive PCR amplification analysis.

As an initial investigation of a potential oncogenic role of CCK-4 in colon carcinoma, polyA$^+$ RNA preparations of fifteen colon carcinoma cell lines were examined for CCK-4 mRNA expression. 9 of the cell lines tested showed elevated levels of CCK-4 transcripts. The levels of CCK-4 mRNA varied considerably, with the strongest expression in SW 480, SW 1414, SW 837, HT 29, SW 620, and CACO2 cell lines. Together with the lack of expression in adult normal colon tissue, this observation suggested that rather than representing a normal housekeeping RTK for colon epithelial cell types, CCK-4 may be related to the pathophysiology of colon carcinomas.

**III. Expression of CCK-4 cDNA in 293 fibroblasts**

To investigate the biochemical properties of the CCK-4 receptor protein, we transfected expression plasmids containing cDNA encoding either CCK-4 or a chimeric receptor, EC-R, consisting of the human EGF-R extracellular domain and the CCK-4 transmembrane and intracellular domains into human 293 embryonic kidney fibroblasts. Cells transfected with expression vectors containing human EGF-R cDNA or the cytomegalovirus early promoter-based vector without cDNA insert were used as a control. Transfected cells were metabolically labeled with [$^{35}$S] methionine and the lysates were subjected to immunoprecipitation with appropriate antibodies. Since no CCK-4-specific antibody was yet available, we added, by PCR-assisted mutagenesis, the influenza virus hemaglutinin (HA) epitope YPYDVPDYA (Pati, 1992), which is recognized by the monoclonal antibody cline 12CA5 (Boehringer and Manheim) to the C-terminus of CCK-4, causing an increase of the calculated molecular weight by 1.13 kD to 119.4kD. The anti-HA antibody recognized a single protein band of apparent $M_r$ 145,000 in lysates of CCK-4 vector-transfected cells, while in plasmid control transfected cells no immunoreactive protein was detected. From lysates of cells transfected with the EC-R chimera expression plasmid, immunoprecipitation with antibodies directed against the extracellular domain of human EGF-R led to the recovery of a protein doublet with apparent $M_r$ of 155,000 and 160,000, while wt EGF-R precipitated under the same conditions migrated at the expected $M_r$ of 175,000. Thus, according to its

mobility in SDS polyacrylamide gel electrophoresis (SDS-PAGE), the apparent MW of CCK-4 is approximately 27 kD higher than the calculated MW, which, in analogy to other cell surface proteins, is likely due to glycosylation of the CCK-4 extracellular domain.

In contrast to EGF-R, CCK-4 and EC-R displayed no detectable levels of tyrosine phosphorylation as determined by immunoblot analysis of immunoprecipitated receptors with mAbs directed against phosphotyrosine. This was true for the EC-R chimera, even after exposure of transfected cells to EGF concentrations up to 300 nanograms per milliliter prior to lysis.

In order to increase the sensitivity of our assay, we also performed *in vitro* kinase experiments. HA-modified CCK-4, EC-R, and EGF-R were immunoprecipitated from transiently transfected 293 cells and either incubated with [g32P]ATP under appropriate reaction conditions, followed by SDS-PAGE and autoradiography or with unlabeled ATP, followed by immunodetection of tyrosine-phosphorylated proteins by antiphosphotyrosine immunoblot analysis. While phosphorylation of EGF-R was readily detectable under both conditions, no phosphorylation was detected for CCK-4 and EC-R, even after prolonged exposure times.

**IV. Nucleic Acid Encoding A CCK-4 Polypeptide.**

Included within the scope of this invention are the functional equivalents of the herein-described isolated nucleic acid molecules. The degeneracy of the genetic code permits substitution of certain codons by

other codons which specify the same amino acid and hence would give rise to the same protein. The nucleic acid sequence can vary substantially since, with the exception of methionine and tryptophan, the known amino acids can be coded for by more than one codon. Thus, portions or all of the CCK-4 gene could be synthesized to give a nucleic acid sequence significantly different from that shown in Figure 1. The encoded amino acid sequence thereof would, however, be preserved.

In addition, the nucleic acid sequence may comprise a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the nucleic acid formula shown in Figure 1 or a derivative thereof. Any nucleotide or polynucleotide may be used in this regard, provided that its addition, deletion or substitution does not alter the amino acid sequence of Figure 1 which is encoded by the nucleotide sequence. For example, the present invention is intended to include any nucleic acid sequence resulting from the addition of ATG as an initiation codon at the 5'-end of the inventive nucleic acid sequence or its derivative, or from the addition of TTA, TAG or TGA as a termination codon at the 3'-end of the inventive nucleotide sequence or its derivative. Moreover, the nucleic acid molecule of the present invention may, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end.

Such functional alterations of a given nucleic acid sequence afford an opportunity to promote secretion

and/or processing of heterologous proteins encoded by foreign nucleic acid sequences fused thereto. All variations of the nucleotide sequence of the CCK-4 genes and fragments thereof permitted by the genetic code are, therefore, included in this invention.

Further, it is possible to delete codons or to substitute one or more codons by codons other than degenerate codons to produce a structurally modified polypeptide, but one which has substantially the same utility or activity of the polypeptide produced by the unmodified nucleic acid molecule. As recognized in the art, the two polypeptides are functionally equivalent, as are the two nucleic acid molecules which give rise to their production, even though the differences between the nucleic acid molecules are not related to degeneracy of the genetic code.

**V. A Nucleic Acid Probe for the Detection of CCK-4.**

A nucleic acid probe of the present invention may be used to probe an appropriate chromosomal or cDNA library by usual hybridization methods to obtain another nucleic acid molecule of the present invention. A chromosomal DNA or cDNA library may be prepared from appropriate cells according to recognized methods in the art (cf. Molecular Cloning: A Laboratory Manual, second edition, edited by Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989).

In the alternative, chemical synthesis is carried out in order to obtain nucleic acid probes having nucleotide sequences which correspond to N-

terminal and C-terminal portions of the amino acid sequence of the polypeptide of interest. Thus, the synthesized nucleic acid probes may be used as primers in a polymerase chain reaction (PCR) carried out in accordance with recognized PCR techniques, essentially according to PCR Protocols, A Guide to Methods and Applications, edited by Michael et al., Academic Press, 1990, utilizing the appropriate chromosomal or cDNA library to obtain the fragment of the present invention.

One skilled in the art can readily design such probes based on the sequence disclosed herein using methods of computer alignment and sequence analysis known in the art (cf. Molecular Cloning: A Laboratory Manual, second edition, edited by Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, 1989). The hybridization probes of the present invention can be labeled by standard labeling techniques such as with a radiolabel, enzyme label, fluorescent label, biotin-avidin label, chemiluminescence, and the like. After hybridization, the probes may be visualized using known methods.

The nucleic acid probes of the present invention include RNA, as well as DNA probes, such probes being generated using techniques known in the art. The nucleic acid probe may be immobilized on a solid support. Examples of such solid supports include, but are not limited to, plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, and acrylic resins, such as polyacrylamide and latex beads.

Techniques for coupling nucleic acid probes to such solid supports are well known in the art.

The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The sample used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample which is compatible with the method utilized.

**VI. A Probe Based Method And Kit For Detecting CCK-4.**

One method of detecting the presence of CCK-4 in a sample comprises a) contacting said sample with the above-described nucleic acid probe, under conditions such that hybridization occurs, and b) detecting the presence of said probe bound to said nucleic acid molecule. One skilled in the art would select the nucleic acid probe according to techniques known in the an as described above. Samples to be tested include but should not be limited to RNA samples of human tissue.

A kit for detecting the presence of CCK-4 in a sample comprises at least one container means having disposed therein the above-described nucleic acid probe. The kit may further comprise other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound nucleic acid

probe. Examples of detection reagents include, but are not limited to radiolabelled probes, enzymatic labeled probes (horse radish peroxidase, alkaline phosphatase), and affinity labeled probes (biotin, avidin, or steptavidin).

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the probe or primers used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and the like), and containers which contain the reagents used to detect the hybridized probe, bound antibody, amplified product, or the like. One skilled in the art will readily recognize that the nucleic acid probes described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

## VII. DNA Constructs Comprising a CCK-4 Nucleic Acid Molecule and Cells Containing These Constructs.

The present invention also relates to a recombinant DNA molecule comprising, 5' to 3', a promoter effective to initiate transcription in a host cell and the above-described nucleic acid molecules. In addition, the present invention relates to a recombinant DNA molecule comprising a vector and an above-described nucleic acid molecules. The present invention also relates to a nucleic acid molecule comprising a transcriptional region functional in a cell, a sequence complimentary to an RNA sequence encoding an amino acid sequence corresponding to the above-described polypeptide, and a transcriptional termination region functional in said cell. The above-described molecules may be isolated and/or purified DNA molecules.

The present invention also relates to a cell or organism that contains an above-described nucleic acid molecule. The peptide may be purified from cells which have been altered to express the peptide. A cell is said to be "altered to express a desired peptide" when the cell, through genetic manipulation, is made to produce a protein which it normally does not produce or which the cell normally produces at lower levels. One skilled in the art can readily adapt procedures for introducing and expressing either genomic, cDNA, or synthetic sequences into either eukaryotic or prokaryotic cells.

A nucleic acid molecule, such as DNA, is said to be "capable of expressing" a polypeptide if it contains nucleotide sequences which contain transcriptional and translational regulatory information

and such sequences are "operably linked" to nucleotide sequences which encode the polypeptide. An operable linkage is a linkage in which the regulatory DNA sequences and the DNA sequence sought to be expressed are connected in such a way as to permit gene sequence expression. The precise nature of the regulatory regions needed for gene sequence expression may vary from organism to organism, but shall in general include a promoter region which, in prokaryotes, contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

If desired, the non-coding region 3' to the sequence encoding an CCK-4 gene may be obtained by the above-described methods. This region may be retained for its transcriptional termination regulatory sequences, such as termination and polyadenylation. Thus, by retaining the 3'-region naturally contiguous to the DNA sequence encoding an CCK-4 gene, the transcriptional termination signals may be provided. Where the transcriptional termination signals are not satisfactorily functional in the expression host cell, then a 3' region functional in the host cell may be substituted.

Two DNA sequences (such as a promoter region sequence and an CCK-4 sequence) are said to be operably

linked if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region sequence to direct the transcription of an CCK-4 gene sequence, or (3) interfere with the ability of the an CCK-4 gene sequence to be transcribed by the promoter region sequence. Thus, a promoter region would be operably linked to a DNA sequence if the promoter were capable of effecting transcription of that DNA sequence. Thus, to express an CCK-4 gene, transcriptional and translational signals recognized by an appropriate host are necessary.

The present invention encompasses the expression of the CCK-4 gene (or a functional derivative thereof) in either prokaryotic or eukaryotic cells. Prokaryotic hosts are, generally, very efficient and convenient for the production of recombinant proteins and are, therefore, one type of preferred expression system for the CCK-4 gene. Prokaryotes most frequently are represented by various strains of E. coli. However, other microbial strains may also be used, including other bacterial strains.

In prokaryotic systems, plasmid vectors that contain replication sites and control sequences derived from a species compatible with the host may be used. Examples of suitable plasmid vectors may include pBR322, pUC118, pUC119 and the like; suitable phage or bacteriophage vectors may include γgt10, γgt11 and the like; and suitable virus vectors may include pMAM-neo, pKRC and the like. Preferably, the selected vector of

the present invention has the capacity to replicate in the selected host cell.

Recognized prokaryotic hosts include bacteria such as E. coil, Bacillus, Streptomyces, Pseudomonas, Salmonella, Serratia, and the like. However, under such conditions, the peptide will not be glycosylated. The prokaryotic host must be compatible with the replicon and control sequences in the expression plasmid.

To express CCK-4 (or a functional derivative thereof) in a prokaryotic cell, it is necessary to operably link the CCK-4 sequence to a functional prokaryotic promoter. Such promoters may be either constitutive or, more preferably, regulatable (i.e., inducible or derepressible). Examples of constitutive promoters include the int promoter of bacteriophage λ, the bla promoter of the β-lactamase gene sequence of pBR322, and the CAT promoter of the chloramphenicol acetyl transferase gene sequence of pPR325, and the like. Examples of inducible prokaryotic promoters include the major right and left promoters of bacteriophage λ ($P_L$ and $P_R$), the trp, recA, lacZ, lacI, and gal promoters of E. coli, the α-amylase (Ulmanen et at., J. Bacteriol. 162:176-182(1985)) and the ς-28-specific promoters of B. subtilis (Gilman et at., Gene sequence 32:11-20(1984)), the promoters of the bacteriophages of Bacillus (Gryczan, In: The Molecular Biology of the Bacilli, Academic Press, Inc., NY (1982)), and Streptomyces promoters (Ward et at., Mol. Gen. Genet. 203:468-478(1986)). Prokaryotic promoters are reviewed by Glick (J. Ind. Microbiot. 1:277-

282(1987)); Cenatiempo (Biochimie 68:505-516(1986)); and Gottesman (Ann. Rev. Genet. 18:415-442 (1984)).

Proper expression in a prokaryotic cell also requires the presence of a ribosome binding site upstream of the gene sequence-encoding sequence. Such ribosome binding sites are disclosed, for example, by Gold et at. (Ann. Rev. Microbiol. 35:365-404(1981)). The selection of control sequences, expression vectors, transformation methods, and the like, are dependent on the type of host cell used to express the gene. As used herein, "cell", "cell line", and "cell culture" may be used interchangeably and all such designations include progeny. Thus, the words "transformants" or "transformed cells" include the primary subject cell and cultures derived therefrom, without regard to the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. However, as defined, mutant progeny have the same functionality as that of the originally transformed cell.

Host cells which may be used in the expression systems of the present invention are not strictly limited, provided that they are suitable for use in the expression of the CCK-4 peptide of interest. Suitable hosts may often include eukaryotic cells. Preferred eukaryotic hosts include, for example, yeast, fungi, insect cells, mammalian cells either in vivo, or in tissue culture. Mammalian cells which may be useful as hosts include HeLa cells, cells of fibroblast origin such as VERO or CHO-K1, or cells of lymphoid origin and

their derivatives. Preferred mammalian host cells include SP2/0 and J558L, as well as neuroblastoma cell lines such as IMR 332 which may provide better capacities for correct post-translational processing.

In addition, plant cells are also available as hosts, and control sequences compatible with plant cells are available, such as the cauliflower mosaic virus 35S and 19S, and nopaline synthase promoter and polyadenylation signal sequences. Another preferred host is an insect cell, for example the Drosophila larvae. Using insect cells as hosts, the Drosophila alcohol dehydrogenase promoter can be used. Rubin, Science 240:1453-1459(1988). Alternatively, baculovirus vectors can be engineered to express large amounts of CCK-4 in insects cells (Jasny, Science 238:1653 (1987); Miller et al., In: Genetic Engineering (1986), Setlow, J.K., et al., eds., Plenum, Vol. 8, pp. 277-297).

Any of a series of yeast gene sequence expression systems can be utilized which incorporate promoter and termination elements from the actively expressed gene sequences coding for glycolytic enzymes are produced in large quantities when yeast are grown in mediums rich in glucose. Known glycolytic gene sequences can also provide very efficient transcriptional control signals. Yeast provides substantial advantages in that it can also carry out post-translational peptide modifications. A number of recombinant DNA strategies exist which utilize strong promoter sequences and high copy number of plasmids which can be utilized for production of the desired proteins in yeast. Yeast

recognizes leader sequences on cloned mammalian gene sequence products and secretes peptides bearing leader sequences (i.e., pre-peptides). For a mammalian host, several possible vector systems are available for the expression of CCK-4.

A wide variety of transcriptional and translational regulatory sequences may be employed, depending upon the nature of the host. The transcriptional and translational regulatory signals may be derived from viral sources, such as adenovirus, bovine papilloma virus, cytomegalovirus, simian virus, or the like, where the regulatory signals are associated with a particular gene sequence which has a high level of expression. Alternatively, promoters from mammalian expression products, such as actin, collagen, myosin, and the like, may be employed. Transcriptional initiation regulatory signals may be selected which allow for repression or activation, so that expression of the gene sequences can be modulated. Of interest are regulatory signals which are temperature-sensitive so that by varying the temperature, expression can be repressed or initiated, or are subject to chemical (such as metabolite) regulation.

Expression of CCK-4 in eukaryotic hosts requires the use of eukaryotic regulatory regions. Such regions will, in general, include a promoter region sufficient to direct the initiation of RNA synthesis. Preferred eukaryotic promoters include, for example, the promoter of the mouse metallothionein I gene sequence (Hamer et al., J. Mol. Appl. Gen. 1:273-288(1982)); the

TK promoter of Herpes virus (McKnight, Cell 31:355-365 (1982)); the SV40 early promoter (Benoist et al., Nature (London) 290:304-310(1981)); the yeast gal4 gene sequence promoter (Johnston et al., Proc. Natl. Acad. Sci. (USA) 79:6971-6975(1982); Silver et al., Proc. Natl. Acad. Sci. (USA) 81:5951-5955 (1984)).

Translation of eukaryotic mRNA is initiated at the codon which encodes the first methionine. For this reason, it is preferable to ensure that the linkage between a eukaryotic promoter and a DNA sequence which encodes CCK-4 (or a functional derivative thereof) does not contain any intervening codons which are capable of encoding a methionine (i.e., AUG). The presence of such codons results either in a formation of a fusion protein (if the AUG codon is in the same reading frame as the CCK-4 coding sequence) or a frame-shift mutation (if the AUG codon is not in the same reading frame as the CCK-4 coding sequence).

A CCK-4 nucleic acid molecule and an operably linked promoter may be introduced into a recipient prokaryotic or eukaryotic cell either as a nonreplicating DNA (or RNA) molecule, which may either be a linear molecule or, more preferably, a closed covalent circular molecule. Since such molecules are incapable of autonomous replication, the expression of the gene may occur through the transient expression of the introduced sequence. Alternatively, permanent expression may occur through the integration of the introduced DNA sequence into the host chromosome.

A vector may be employed which is capable of integrating the desired gene sequences into the host cell chromosome. Cells which have stably integrated the introduced DNA into their chromosomes can be selected by also introducing one or more markers which allow for selection of host cells which contain the expression vector. The marker may provide for prototrophy to an auxotrophic host, biocide resistance, e.g., antibiotics, or heavy metals, such as copper, or the like. The selectable marker gene sequence can either be directly linked to the DNA gene sequences to be expressed, or introduced into the same cell by co-transfection. Additional elements may also be needed for optimal synthesis of single chain binding protein mRNA. These elements may include splice signals, as well as transcription promoters, enhancers, and termination signals. cDNA expression vectors incorporating such elements include those described by Okayama, Molec. Cell. Biol. 3:280(1983).

The introduced nucleic acid molecule can be incorporated into a plasmid or viral vector capable of autonomous replication in the recipient host. Any of a wide variety of vectors may be employed for this purpose. Factors of importance in selecting a particular plasmid or viral vector include: the ease with which recipient cells that contain the vector may be recognized and selected from those recipient cells which do not contain the vector; the number of copies of the vector which are desired in a particular host; and whether it is desirable to be able to "shuttle" the

vector between host cells of different species. Preferred prokaryotic vectors include plasmids such as those capable of replication in E. coil (such as, for example, pBR322, ColEl, pSC101, pACYC 184, πVX. Such plasmids are, for example, disclosed by Sambrook (cf. Molecular Cloning: A Laboratory Manual, second edition, edited by Sambrook, Fritsch, & Maniatis, Cold Spring Harbor Laboratory, (1989)). Bacillus plasmids include pC194, pC221, pT127, and the like. Such plasmids are disclosed by Gryczan (In: The Molecular Biology of the Bacilli, Academic Press, NY (1982), pp. 307-329). Suitable Streptomyces plasmids include p1J101 (Kendall et al., J. Bacteriol. 169:4177-4183 (1987)), and streptomyces bacteriophages such as ϕC31 (Chater et al., In: Sixth International Symposium on Actinomycetales Biology, Akademiai Kaido, Budapest, Hungary (1986), pp. 45-54). Pseudomonas plasmids are reviewed by John et al. (Rev. Infect. Dis. 8:693-704(1986)), and Izaki (Jpn. J. Bacteriol. 33:729-742(1978)).

Preferred eukaryotic plasmids include, for example, BPV, vaccinia, SV40, 2-micron circle, and the like, or their derivatives. Such plasmids are well known in the art (Botstein et al., Miami Wntr. Symp. 19:265-274(1982); Broach, In: The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, p. 445-470 (1981); Broach, Cell 28:203-204 (1982); Bollon et at., J. Ctin. Hematol. Oncol. 10:39-48 (1980); Maniatis, In: Cell Biology: A Comprehensive Treatise,

Vol. 3, Gene Sequence Expression, Academic Press, NY, pp. 563-608(1980).

Once the vector or nucleic acid molecule containing the construct(s) has been prepared for expression, the DNA construct(s) may be introduced into an appropriate host cell by any of a variety of suitable means, i.e., transformation, transfection, conjugation, protoplast fusion, electroporation, particle gun technology, calcium phosphate-precipitation, direct microinjection, and the like. After the introduction of the vector, recipient cells are grown in a selective medium, which selects for the growth of vector-containing cells. Expression of the cloned gene molecule(s) results in the production of CCK-4 or fragments thereof. This can take place in the transformed cells as such, or following the induction of these cells to differentiate (for example, by administration of bromodeoxyuracil to neuroblastoma cells or the like). A variety of incubation conditions can be used to form the peptide of the present invention. The most preferred conditions are those which mimic physiological conditions.

**VIII. Purified CCK-4 Polypeptides**

A variety of methodologies known in the art can be utilized to obtain the peptide of the present invention. The peptide may be purified from tissues or cells which naturally produce the peptide. Alternatively, the above-described isolated nucleic acid fragments could be used to expressed the CCK-4 protein

in any organism. The samples of the present invention include cells, protein extracts or membrane extracts of cells, or biological fluids. The sample will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts used as the sample.

Any eukaryotic organism can be used as a source for the peptide of the invention, as long as the source organism naturally contains such a peptide. As used herein, "source organism" refers to the original organism from which the amino acid sequence of the subunit is derived, regardless of the organism the subunit is expressed in and ultimately isolated from.

One skilled in the art can readily follow known methods for isolating proteins in order to obtain the peptide free of natural contaminants. These include, but are not limited to: size-exclusion chromatography, HPLC, ion-exchange chromatography, and immuno-affinity chromatography.

**IX. An Antibody Having Binding Affinity To A CCK-4 Polypeptide And A Hybridoma Containing the Antibody.**

The present invention relates to an antibody having binding affinity to a CCK-4 polypeptide. The polypeptide may have the amino acid sequence set forth in Figure 1, or functional derivitive thereof, or at least 9 contiguous amino acids thereof (preferably, at least 20, 30, 35, or 40 contiguous amino acids thereof).

The present invention also relates to an antibody having specific binding affinity to an CCK-4

polypeptide. Such an antibody may be isolated by comparing its binding affinity to a CCK-4 polypeptide with its binding affinity to another polypeptide. Those which bind selectively to CCK-4 would be chosen for use in methods requiring a distinction between CCK-4 and other polypeptides. Such methods could include, but should not be limited to, the analysis of altered CCK-4 expression in tissue containing other polypeptides such as FAK.

The CCK-4 proteins of the present invention can be used in a variety of procedures and methods, such as for the generation of antibodies, for use in identifying pharmaceutical compositions, and for studying DNA/protein interaction.

The CCK-4 peptide of the present invention can be used to produce antibodies or hybridomas. One skilled in the art will recognize that if an antibody is desired, such a peptide would be generated as described herein and used as an immunogen. The antibodies of the present invention include monoclonal and polyclonal antibodies, as well fragments of these antibodies, and humanized forms. Humanized forms of the antibodies of the present invention may be generated using one of the procedures known in the art such as chimerization or CDR grafting. The present invention also relates to a hybridoma which produces the above-described monoclonal antibody, or binding fragment thereof. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing monoclonal antibodies and hybridomas are well known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth et al., J. Immunol. Methods 35:1-21(1980)). Any animal (mouse, rabbit, and the like) which is known to produce antibodies can be immunized with the selected polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of polypeptide used for immunization will vary based on the animal which is immunized, the antigenicity of the polypeptide and the site of injection.

The polypeptide may be modified or administered in an adjuvant in order to increase the peptide antigenicity. Methods of increasing the antigenicity of a polypeptide are well known in the art. Such procedures include coupling the antigen with a heterologous protein (such as globulin or β-galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Ag14 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells. Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an

antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz et al., Exp. Cell Res. 175:109-124(1988)). Hybridomas secreting the desired antibodies are cloned and the class and subclass is determined using procedures known in the art (Campbell, Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology, supra (1984)).

For polyclonal antibodies, antibody containing antisera is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures. The above-described antibodies may be detectably labeled. Antibodies can be detectably labeled through the use of radioisotopes, affinity labels (such as biotin, avidin, and the like), enzymatic labels (such as horse radish peroxidase, alkaline phosphatase, and the like) fluorescent labels (such as FITC or rhodamine, and the like), paramagnetic atoms, and the like. Procedures for accomplishing such labeling are well-known in the art, for example, see (Stemberger et al., J. Histochem. Cytochem. 18:315(1970); Bayer et at., Meth. Enzym. 62:308(1979); Engval et al., Immunot. 109:129(1972); Goding, J. Immunol. Meth. 13:215(1976)). The labeled antibodies of the present invention can be used for in vitro, in vivo, and in situ assays to identify cells or tissues which express a specific peptide.

The above-described antibodies may also be immobilized on a solid support. Examples of such solid

supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10(1986); Jacoby et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for in vitro, in vivo, and in situ assays as well as in immunochromotography.

Furthermore, one skilled in the art can readily adapt currently available procedures, as well as the techniques, methods and kits disclosed above with regard to antibodies, to generate peptides capable of binding to a specific peptide sequence in order to generate rationally designed antipeptide peptides, for example see Hurby et al., "Application of Synthetic Peptides: Antisense Peptides", In Synthetic Peptides, A User's Guide, W.H. Freeman, NY, pp. 289-307(1992), and Kaspczak et al., Biochemistry 28:9230-8(1989).

Anti-peptide peptides can be generated by replacing the basic amino acid residues found in the CCK-4 peptide sequence with acidic residues, while maintaining hydrophobic and uncharged polar groups. For example, lysine, arginine, and/or histidine residues are replaced with aspartic acid or glutamic acid and glutamic acid residues are replaced by lysine, arginine or histidine.

**X. An Antibody Based Method And Kit For Detecting CCK-4.**

The present invention encompasses a method of detecting an CCK-4 polypeptide in a sample, comprising: a) contacting the sample with an above-described antibody, under conditions such that immunocomplexes form, and b) detecting the presence of said antibody bound to the polypeptide. In detail, the methods comprise incubating a test sample with one or more of the antibodies of the present invention and assaying whether the antibody binds to the test sample. Altered levels of CCK-4 in a sample as compared to normal levels may indicate muscular disease.

Conditions for incubating an antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the antibody used in the assay. One skilled in the art will recognize that any one of the commonly available immunological assay formats (such as radioimmunoassays, enzyme-linked immunosorbent assays, diffusion based Ouchterlony, or rocket immunofluorescent assays) can readily be adapted to employ the antibodies of the present invention. Examples of such assays can be found in Chard, "An Introduction to Radioimmunoassay and Related Techniques" Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock et al., "Techniques in Immunocytochemistry," Academic Press, Orlando, FL Vol. 1(1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, "Practice and Theory of Enzyme Immunoassays: Laboratory

Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands (1985).

The immunological assay test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is capable with the system utilized.

A kit contains all the necessary reagents to carry out the previously described methods of detection. The kit may comprise: i) a first container means containing an above-described antibody, and ii) second container means containing a conjugate comprising a binding partner of the antibody and a label. In another preferred embodiment, the kit further comprises one or more other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound antibodies.

Examples of detection reagents include, but are not limited to, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the chromophoric, enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. The compartmentalized kit may be as described above for

nucleic acid probe kits. One skilled in the art will readily recognize that the antibodies described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

**XI. Isolation of Compounds Which Interact With CCK-4.**

The present invention also relates to a method of detecting a compound capable of binding to a CCK-4 polypeptide comprising incubating the compound with CCK-4 and detecting the presence of the compound bound to CCK-4. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts.

The present invention also relates to a method of detecting an agonist or antagonist of CCK-4 activity comprising incubating cells that produce CCK-4 in the presence of a compound and detecting changes in the level of CCK-4 activity. The compounds thus identified would produce a change in activity indicative of the presence of the compound. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts. Once the compound is identified it can be isolated using techniques well known in the art.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing CCK-4 associated activity in a mammal comprising administering to said mammal an agonist or antagonist to CCK-4 in an

amount sufficient to effect said agonism or antagonism. A method of treating diabetes mellitus, skeletal muscle diseases, Alzheimer's disease, or peripheral neuropathies in a mammal with an agonist or antagonist of CCK-4 activity comprising administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize CCK-4 associated functions is also encompassed in the present application.

**XII. Transgenic Animals.**

A variety of methods are available for the production of transgenic animals associated with this invention. DNA can be injected into the pronucleus of a fertilized egg before fusion of the male and female pronuclei, or injected into the nucleus of an embryonic cell (e.g., the nucleus of a two-cell embryo) following the initiation of cell division (Brinster et al., Proc. Nat. Acad. Sci. USA 82: 4438-4442 (1985)). Embryos can be infected with viruses, especially retroviruses, modified to carry inorganic-ion receptor nucleotide sequences of the invention.

Pluripotent stem cells derived from the inner cell mass of the embryo and stabilized in culture can be manipulated in culture to incorporate nucleotide sequences of the invention. A transgenic animal can be produced from such cells through implantation into a blastocyst that is implanted into a foster mother and allowed to come to term. Animals suitable for transgenic experiments can be obtained from standard commercial sources such as Charles River (Wilmington,

MA), Taconic (Germantown, NY), Harlan Sprague Dawley (Indianapolis, IN), etc.

The procedures for manipulation of the rodent embryo and for microinjection of DNA into the pronucleus of the zygote are well known to those of ordinary skill in the art (Hogan et al., supra). Microinjection procedures for fish, amphibian eggs and birds are detailed in Houdebine and Chourrout, Experientia 47: 897-905 (1991). Other procedures for introduction of DNA into tissues of animals are described in U.S. Patent No., 4,945,050 (Sandford et al., July 30, 1990).

By way of example only, to prepare a transgenic mouse, female mice are induced to superovulate. Females are placed with males, and the mated females are sacrificed by $CO_2$ asphyxiation or cervical dislocation and embryos are recovered from excised oviducts. Surrounding cumulus cells are removed. Pronuclear embryos are then washed and stored until the time of injection. Randomly cycling adult female mice are paired with vasectomized males. Recipient females are mated at the same time as donor females. Embryos then are transferred surgically. The procedure for generating transgenic rats is similar to that of mice. See Hammer et al., Cell 63:1099-1112 (1990).

Methods for the culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection also are well known

to those of ordinary skill in the art. See, for example, Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E.J. Robertson, ed., IRL Press (1987).

In cases involving random gene integration, a clone containing the sequence(s) of the invention is co-transfected with a gene encoding resistance. Alternatively, the gene encoding neomycin resistance is physically linked to the sequence(s) of the invention. Transfection and isolation of desired clones are carried out by any one of several methods well known to those of ordinary skill in the art (E.J. Robertson, supra).

DNA molecules introduced into ES cells can also be integrated into the chromosome through the process of homologous recombination. Capecchi, Science 244: 1288-1292 (1989). Methods for positive selection of the recombination event (i.e., neo resistance) and dual positive-negative selection (i.e., neo resistance and gancyclovir resistance) and the subsequent identification of the desired clones by PCR have been described by Capecchi, supra and Joyner et al., Nature 338: 153-156 (1989), the teachings of which are incorporated herein. The final phase of the procedure is to inject targeted ES cells into blastocysts and to transfer the blastocysts into pseudopregnant females. The resulting chimeric animals are bred and the offspring are analyzed by Southern blotting to identify individuals that carry the transgene. Procedures for the production of non-rodent mammals and other animals have been discussed by others. See Houdebine and

Chourrout, supra; Pursel et al., Science 244:1281-1288 (1989); and Simms et al., Bio/Technology 6:179-183 (1988).

**XIII. Compositions**

The present invention relates to removing or reducing an abnormality in a signal transduction pathway, wherein the signal transduction pathway contains a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner. The present invention also relates to compositions and methods for the treatment of disorders which involve modulating the activity and/or level of individual components, and relates to methods for the identification of agents for such treatments. Additionally, the present invention relates to methods and compositions for prognostic evaluation of such disorders.

Described herein are compositions and methods for the prevention, prognostic evaluation, and treatment of disorders in which a CCK-4 receptor tyrosine kinase may be involved, in particular, cell proliferative disorders, especially cancer, in which a CCK-4 receptor tyrosine kinase is involved.

First, methods and compositions for the treatment of such disorders are described. Such methods and compositions may include, but are not limited to the agents capable of decreasing or inhibiting the interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner and agents capable of inhibiting or decreasing the activity of such complexes, agents capable of modulating the activity and/or level of indi-

vidual components of the proteins, and the use and administration of such agents. Agents capable of modulating the activity and/or level of interaction between CCK-4 receptor tyrosine kinase and a CCK-4 binding partner include those agents that inhibit or decrease the dephosphorylating activity of tyrosine phosphatases.

Second, methods are described for the identification of such agents. These methods may include, for example, assays to identify agents capable of disrupting or inhibiting or promoting the interaction between components of the complexes (e.g., CCK-4:binding partner complexes), and may also include paradigms and strategies for the rational design of drugs capable of disruption and/or inhibition and/or promotion of such complexes.

**XIV. Binding partner/Receptor Complexes**

The complexes involved in the invention include a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner or derivatives thereof, as described below. Under standard physiological conditions, the components of such complexes are capable of forming stable, non-covalent attachments with one or more of the other complex components. Methods for the purification and production of such protein complexes, and of cells that exhibit such complexes are described below.

The complexes involved in the invention also include tyrosine phosphatases responsible for dephosphorylating activated CCK-4 receptors, thus modulating the ability to bind to a binding partner and other signal

transduction components. Identification of such tyrosine phosphatase(s) may be accomplished using techniques known to one skilled in the art.

**XV. Disruption of Protein Complexes**

Disruption of complexes (e.g., CCK-4:binding partner complexes), for example by decreasing or inhibiting or promoting the interactions between component members of such a complex may have differing modulatory effects on the event involved, depending on the individual protein complex. "Disruption", as used here, is meant to refer not only to a physical separation of protein complex components, but also refers to a perturbation of the activity of the complexes, regardless of whether or not such complexes remain able, physically, to form. "Activity", as used here, refers to the function of the protein complex in the signal transduction cascade of the cell in which such a complex is formed, i.e., refers to the function of the complex in effecting or inhibiting a transduction of an extracellular signal into a cell. For example, the effect of complex disruption may augment, reduce, or block a signal normally transduced into the cell. Likewise, depending on the disorder involved, either augmentation, reduction, or blockage of a signal normally transduced into the cell will be desirable for the treatment of the disorder.

A disorder involving a complex may, for example, develop because the presence of such a complex brings about the aberrant inhibition of a normal signal transduction event. In such a case, the disruption of

the complex would allow the restoration of the usual signal transduction event. Further, an aberrant complex may bring about an altered subcellular adapter protein localization, which may result in, for example, dysfunctional cellular events. An inhibition of the complex in this case would allow for restoration or maintenance of a normal cellular architecture. Still further, an agent or agents that cause(s) disruption of the complex may bring about the disruption of the interactions among other potential components of a complex.

Nucleotide sequences encoding peptide agents which are to be utilized intracellularly may be expressed in the cells of interest, using techniques which are well known to those of ordinary skill in the art. For example, expression vectors derived from viruses such as retroviruses, vaccinia virus, adenoviruses, adeno-associated virus, herpes viruses, or bovine papilloma virus, may be used for delivery and expression of such nucleotide sequences into the targeted cell population. Methods for the construction of such vectors are well known. See, for example, the techniques described in Maniatis et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y. and in Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, N.Y, 1989. Complex-binding domains can be identified using, for example, techniques such as those described in Rotin et al. (Rotin et al., *EMBO J.* 11:559-567, 1992), Songyang et al. (Songyang et al., *Cell* 72:767-778, 1993), Felder et al., *Mol. Cell. Biol.*

13:1449-1455, 1993), Fantl et al. (*Cell* 69:413-422, 1992), and Domchek et al. (*Biochemistry* 31:9865-9870, 1992).

Alternatively, antibodies capable of interfering with complex formation may be produced as described below and administered for the treatment of disorders involving a component capable of forming a complex with another protein. For example, neutralizing antibodies which are capable of interfering with ligand binding may be administered using standard techniques. Alternatively, nucleotide sequences encoding single-chain antibodies may be expressed within the target cell population by utilizing, for example, techniques such as those described in Marasco et al. (Marasco et al., *Proc. Natl. Acad. Sci. USA* 90:7889-7893, 1993).

Agents which act intracellularly to interfere with the formation and/or activity of the protein complexes of the invention may also be small organic or inorganic compounds. A method for identifying these and other intracellular agents is described below.

**XVI. Antibodies to Complexes**

Described herein are methods for the production of antibodies which are capable of specifically recognizing a complex or an epitope thereof, or of specifically recognizing an epitope on either of the components of the complex, especially those epitopes which would not be recognized by the antibody when the component is present separate and apart from the complex. Such antibodies may include, but are not limited

to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, fragments produced by a FAb expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a complex in a biological sample, or, alternatively, as a method for the inhibition of a complex formation, thus inhibiting the development of a disorder.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as a complex, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, various host animals may be immunized by injection with the complex including but not limited to rabbits, mice, rats, etc. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

A monoclonal antibody, which is a substantially homogeneous population of antibodies to a particular antigen, may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but

are not limited to the hybridoma technique of Kohler and Milstein (*Nature* 256:495-497, 1975) and U.S. Patent No. 4,376,110), the human B-cell hybridoma technique (Kosbor et al., *Immunology Today* 4:72, 1983; Cole et al., *Proc. Natl. Acad. Sci. USA* 80:2026-2030, 1983), and the EBV-hybridoma technique (Cole et al., Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., 1985, pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., *Proc. Natl. Acad. Sci.*, 81:6851-6855, 1984; Neuberger et al., *Nature*, 312:604-608, 1984; Takeda et al., *Nature*, 314:452-454, 1985) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent 4,946,778; Bird, *Science* 242:423-426, 1988; Huston et al., *Proc. Natl. Acad. Sci. USA* 85:5879-5883, 1988; and Ward et al., *Nature* 334:544-546, 1989) can be

adapted to produce complex-specific single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragment of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which contain specific binding sites of a complex may be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity to the PTK/adapter complex.

One or more components of a protein complex may be present at a higher than normal cellular level (i.e., higher than the concentration known to usually be present in the cell type exhibiting the protein complex of interest) and/or may exhibit an abnormally increased level of cellular activity (i.e., greater than the activity known to usually be present in the cell type exhibiting the protein complex of interest).

For example, the gene encoding a protein complex component may begin to be overexpressed, or may be amplified (i.e., its gene copy number may be increased) in certain cells, leading to an increased number of component molecules within these cells. Additionally, a

gene encoding a protein complex component may begin to express a modified protein product that exhibits a greater than normal level of activity. "Activity", here, refers to the normal cellular function of the component, either enzymatic or structural whose function may include, for example, bringing two or more cellular molecules into the appropriate proximity.

Such an increase in the cellular level and/or activity of a protein complex may lead to the development of a disorder. Treatment of such disorders may, therefore, be effectuated by the administration of agents which decrease the cellular level and/or the activity of the overexpressed and/or overactive protein complex component.

Techniques for decreasing the cellular level and/or the activity of one or more of the protein complex components of interest may include, but are not limited to antisense or ribozyme approaches, and/or gene therapy approaches, each of which is well known to those of skill in the art.

**XVII. Antisense and Ribozyme Approaches to Provide or Disrupt the Complexes of the Present Invention**

Included in the scope of the invention are oligoribonucleotides, including antisense RNA and DNA molecules and ribozymes that function to inhibit translation of one or more components of a protein complex. Anti-sense RNA and DNA molecules act to directly block the translation of mRNA by binding to targeted mRNA and preventing protein translation. With respect to anti-

sense DNA, oligodeoxyribonucleotides derived from the translation initiation site, e.g., between -10 and +10 regions of the relevant nucleotide sequence, are preferred.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific interaction of the ribozyme molecule to complementary target RNA, followed by a endonucleolytic cleavage. Within the scope of the invention are engineered hammerhead or other motif ribozyme molecules that specifically and efficiently catalyze endonucleolytic cleavage of RNA sequences encoding protein complex components.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features, such as secondary structure, that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridization with complementary oligonucleotides, using ribonuclease protection assays. See, Draper PCT WO 93/23569.

Both anti-sense RNA and DNA molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of RNA molecules.

See, Draper, id. hereby incorporated by reference herein. These include techniques for chemically synthesizing oligodeoxyribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Various modifications to the DNA molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribo- or deoxy- nucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

**Gene Therapy**

CCK-4 or its genetic sequences will also be useful in gene therapy (reviewed in Miller, *Nature* 357:455-460, (1992). Miller states that advances have resulted in practical approaches to human gene therapy that have demonstrated positive initial results. The

basic science of gene therapy is described in Mulligan, *Science* 260:926-931, (1993).

In one preferred embodiment, an expression vector containing the CCK-4 coding sequence is inserted into cells, the cells are grown *in vitro* and then infused in large numbers into patients. In another preferred embodiment, a DNA segment containing a promoter of choice (for example a strong promoter) is transferred into cells containing an endogenous CCK-4 in such a manner that the promoter segment enhances expression of the endogenous CCK-4 gene (for example, the promoter segment is transferred to the cell such that it becomes directly linked to the endogenous CCK-4 gene).

The gene therapy may involve the use of an adenovirus containing CCK-4 cDNA targeted to a tumor, systemic CCK-4 increase by implantation of engineered cells, injection with CCK-4 virus, or injection of naked CCK-4 DNA into appropriate tissues.

Target cell populations may be modified by introducing altered forms of one or more components of the protein complexes in order to modulate the activity of such complexes. For example, by reducing or inhibiting a complex component activity within target cells, an abnormal signal transduction event(s) leading to a condition may be decreased, inhibited, or reversed. Deletion or missense mutants of a component, that retain the ability to interact with other components of the protein complexes but cannot function in signal trans-

duction may be used to inhibit an abnormal, deleterious signal transduction event.

Expression vectors derived from viruses such as retroviruses, vaccinia virus, adenovirus, adeno-associated virus, herpes viruses, several RNA viruses, or bovine papilloma virus, may be used for delivery of nucleotide sequences (e.g., cDNA) encoding recombinant CCK-4 protein into the targeted cell population (e.g., tumor cells). Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors containing coding sequences. See, for example, the techniques described in Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y. (1989), and in Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, N.Y. (1989). Alternatively, recombinant nucleic acid molecules encoding protein sequences can be used as naked DNA or in reconstituted system e.g., liposomes or other lipid systems for delivery to target cells (See e.g., Felgner et al., *Nature* 337:387-8, 1989). Several other methods for the direct transfer of plasmid DNA into cells exist for use in human gene therapy and involve targeting the DNA to receptors on cells by complexing the plasmid DNA to proteins. See, Miller, supra.

In its simplest form, gene transfer can be performed by simply injecting minute amounts of DNA into the nucleus of a cell, through a process of microinjection. Capecchi MR, Cell 22:479-88 (1980). Once recombinant genes are introduced into a cell, they

can be recognized by the cells normal mechanisms for transcription and translation, and a gene product will be expressed. Other methods have also been attempted for introducing DNA into larger numbers of cells. These methods include: transfection, wherein DNA is precipitated with $CaPO_4$ and taken into cells by pinocytosis (Chen C. and Okayama H, Mol. Cell Biol. 7:2745-52 (1987)); electroporation, wherein cells are exposed to large voltage pulses to introduce holes into the membrane (Chu G. et al., Nucleic Acids Res., 15:1311-26 (1987)); lipofection/liposome fusion, wherein DNA is packaged into lipophilic vesicles which fuse with a target cell (Felgner PL., et al., Proc. Natl. Acad. Sci. USA. 84:7413-7 (1987)); and particle bombardment using DNA bound to small projectiles (Yang NS. et al., Proc. Natl. Acad. Sci. 87:9568-72 (1990)). Another method for introducing DNA into cells is to couple the DNA to chemically modified proteins.

It has also been shown that adenovirus proteins are capable of destabilizing endosomes and enhancing the uptake of DNA into cells. The admixture of adenovirus to solutions containing DNA complexes, or the binding of DNA to polylysine covalently attached to adenovirus using protein crosslinking agents substantially improves the uptake and expression of the recombinant gene. Curiel DT et al., Am. J. Respir. Cell. Mol. Biol., 6:247-52 (1992).

As used herein "gene transfer" means the process of introducing a foreign nucleic acid molecule into a cell. Gene transfer is commonly performed to

enable the expression of a particular product encoded by the gene. The product may include a protein, polypeptide, anti-sense DNA or RNA, or enzymatically active RNA. Gene transfer can be performed in cultured cells or by direct administration into animals. Generally gene transfer involves the process of nucleic acid contact with a target cell by non-specific or receptor mediated interactions, uptake of nucleic acid into the cell through the membrane or by endocytosis, and release of nucleic acid into the cytoplasm from the plasma membrane or endosome. Expression may require, in addition, movement of the nucleic acid into the nucleus of the cell and binding to appropriate nuclear factors for transcription.

As used herein "gene therapy" is a form of gene transfer and is included within the definition of gene transfer as used herein and specifically refers to gene transfer to express a therapeutic product from a cell *in vivo* or *in vitro*. Gene transfer can be performed *ex vivo* on cells which are then transplanted into a patient, or can be performed by direct administration of the nucleic acid or nucleic acid-protein complex into the patient.

In another preferred embodiment, a vector having nucleic acid sequences encoding CCK-4 is provided in which the nucleic acid sequence is expressed only in specific tissue. Methods of achieving tissue-specific gene expression as set forth in International Publication No. WO 93/09236, filed November 3, 1992 and published May 13, 1993.

In all of the preceding vectors set forth above, a further aspect of the invention is that the nucleic acid sequence contained in the vector may include additions, deletions or modifications to some or all of the sequence of the nucleic acid, as defined above.

In another preferred embodiment, a method of gene replacement is set forth. "Gene replacement" as used herein means supplying a nucleic acid sequence which is capable of being expressed *in vivo* in an animal and thereby providing or augmenting the function of an endogenous gene which is missing or defective in the animal.

**XVIII. Pharmaceutical Formulations and Modes of Administration**

The particular compound, antibody, antisense or ribozyme molecule that affects the protein complexes and the disorder of interest can be administered to a patient either by themselves, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s).

In treating a patient exhibiting an oncogenic disorder of interest, a therapeutically effective amount of a agent or agents such as these is administered. A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical

procedures in cell cultures or experimental animals, e.g., for determining the $LD_{50}$ (the dose lethal to 50% of the population) and the $ED_{50}$ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio $LD_{50}/ED_{50}$. Compounds which exhibit large therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the $ED_{50}$ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized.

For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the $IC_{50}$ as determined in cell culture (i.e., the concentration of the test compound which achieves a half-maximal disruption of the protein complex, or a half-maximal inhibition of the cellular level and/or activity of a complex component). Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by HPLC.

The exact formulation, route of administration and dosage can be chosen by the individual physician in

view of the patient's condition. (See e.g. Fingl et al., in The Pharmacological Basis of Therapeutics, 1975, Ch. 1 p. 1).

It should be noted that the attending physician would know how to and when to terminate, interrupt, or adjust administration due to toxicity, or to organ dysfunctions. Conversely, the attending physician would also know to adjust treatment to higher levels if the clinical response were not adequate (precluding toxicity). The magnitude of an administrated dose in the management of the oncogenic disorder of interest will vary with the severity of the condition to be treated and to the route of administration. The severity of the condition may, for example, be evaluated, in part, by standard prognostic evaluation methods. Further, the dose and perhaps dose frequency, will also vary according to the age, body weight, and response of the individual patient. A program comparable to that discussed above may be used in veterinary medicine.

Depending on the specific conditions being treated, such agents may be formulated and administered systemically or locally. Techniques for formulation and administration may be found in Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co., Easton, PA (1990). Suitable routes may include oral, rectal, transdermal, vaginal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous,

intraperitoneal, intranasal, or intraocular injections, just to name a few.

For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For such transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

Use of pharmaceutically acceptable carriers to formulate the compounds herein disclosed for the practice of the invention into dosages suitable for systemic administration is within the scope of the invention. With proper choice of carrier and suitable manufacturing practice, the compositions of the present invention, in particular, those formulated as solutions, may be administered parenterally, such as by intravenous injection. The compounds can be formulated readily using pharmaceutically acceptable carriers well known in the art into dosages suitable for oral administration. Such carriers enable the compounds of the invention to be formulated as tablets, pills, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated.

Agents intended to be administered intracellularly may be administered using techniques well known to those of ordinary skill in the art. For example, such agents may be encapsulated into liposomes, then administered as described above. Liposomes are spherical lipid

bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the aqueous interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are efficiently delivered into the cell cytoplasm. Additionally, due to their hydrophobicity, small organic molecules may be directly administered intracellularly.

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. The preparations formulated for oral administration may be in the form of tablets, dragees, capsules, or solutions.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Pharmaceutical preparations for oral use can be obtained by combining the active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added.

The nucleic acid sequence encoding CCK-4 can be administered prophylactically, or to patients having a disorder listed above, *e.g.*, by exogenous delivery of the nucleic acid sequence encoding CCK-4 as naked DNA, DNA associated with specific carriers, or in a nucleic acid expression vector to a desired tissue by means of an appropriate delivery vehicle, *e.g.*, a liposome, by use of iontophoresis, electroporation and other pharmacologically approved methods of delivery. Routes

of administration may include intramuscular, intravenous, aerosol, oral (tablet or pill form), topical, systemic, ocular, as a suppository, intraperitoneal and/or intrathecal.

Some methods of delivery that may be used include:

a. encapsulation in liposomes,

b. transduction by retroviral vectors,

c. localization to nuclear compartment utilizing nuclear targeting site found on most nuclear proteins,

d. transfection of cells *ex vivo* with subsequent reimplantation or administration of the transfected cells,

e. a DNA transporter system.

A CCK-4 nucleic acid sequence may be administered utilizing an *ex vivo* approach whereby cells are removed from an animal, transduced with the CCK-4 nucleic acid sequence and reimplanted into the animal. The liver can be accessed by an *ex vivo* approach by removing hepatocytes from an animal, transducing the hepatocytes *in vitro* with the CCK-4 nucleic acid sequence and reimplanting them into the animal (*e.g.*, as described for rabbits by Chowdhury et al, Science 254: 1802-1805, 1991, or in humans by Wilson, Hum. Gene Ther. 3: 179-222, 1992) incorporated herein by reference.

Many nonviral techniques for the delivery of a CCK-4 nucleic acid sequence into a cell can be used, including direct naked DNA uptake (*e.g.*, Wolff et al., Science 247: 1465-1468, 1990), receptor-mediated DNA

uptake, *e.g.*, using DNA coupled to asialoorosomucoid which is taken up by the asialoglycoprotein receptor in the liver (Wu and Wu, J. Biol. Chem. 262: 4429-4432, 1987; Wu et al., J. Biol. Chem. 266: 14338-14342, 1991), and liposome-mediated delivery (*e.g.*, Kaneda et al., Expt. Cell Res. 173: 56-69, 1987; Kaneda et al., Science 243: 375-378, 1989; Zhu et al., Science 261: 209-211, 1993). Many of these physical methods can be combined with one another and with viral techniques; enhancement of receptor-mediated DNA uptake can be effected, for example, by combining its use with adenovirus (Curiel et al., Proc. Natl. Acad. Sci. USA 88: 8850-8854, 1991; Cristiano et al., Proc. Natl. Acad. Sci. USA 90: 2122-2126, 1993).

The CCK-4 or nucleic acid encoding CCK-4 may also be administered via an implanted device that provides a support for growing cells. Thus, the cells may remain in the implanted device and still provide the useful and therapeutic agents of the present invention.

**XIX. Identification of Agents**

The complexes, components of such complexes, functional equivalents thereof, and/or cell lines that express such components and exhibit such protein complexes may be used to screen for additional compounds, antibodies, or other molecules capable of modulating the signal transduction event such complexes are involved in. Methods for purifying and/or producing such complexes, components of the complexes, functional equivalents thereof, and/or cell lines are described herein.

The compounds, antibodies, or other molecules identified may, for example, act to disrupt the protein complexes of the invention (i.e., decrease or inhibit interactions between component members of the complexes, thereby causing physical separation of the components, and/or perturbing the activity of the complexes) or may lower the cellular level and/or decrease the activity of one or more of the components of such complexes.

Such compounds may include, but are not limited to, peptides made of D- and/or L-configuration amino acids (in, for example, the form of random peptide libraries; see Lam et al., *Nature* 354:82-84, 1991), phosphopeptides (in, for example, the form of random or partially degenerate, directed phosphopeptide libraries, see Songyang et al., *Cell* 767-778, 1993), antibodies, and small organic or inorganic molecules. Synthetic compounds, natural products, and other sources of potentially biologically active materials may be screened in a variety of ways, as described herein. The compounds, antibodies, or other molecules identified may be used as oncogenic disorder treatments, as described herein.

Compounds that bind to individual components, or functional portions of the individual components of the complexes (and may additionally be capable of disrupting complex formation) may be identified.

One such method included within the scope of the invention is a method for identifying an agent to be tested for an ability to modulate a signal transduction pathway disorder. The method involves exposing at least

one agent to a protein comprising a functional portion of a member of the protein complex for a time sufficient to allow binding of the agent to the functional portion of the member; removing non-bound agents; and determining the presence of the compound bound to the functional portion of the member of the protein complex, thereby identifying an agent to be tested for an ability to modulate a disorder involving a polypeptide complex.

By "signal transduction disorder" is meant any disease or condition associated with an abnormality in a signal transduction pathway. The protein complex referred to below is a physical association of a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner. The level of interaction between the two components of the complex may be abnormal and thus cause the abnormality in the signal transduction pathway. Alternatively, the level of interaction between the complex components may be normal, but affecting that interaction may effectively treat a signal transduction pathway disorder.

The term "protein" refers to a compound formed of 5-50 or more amino acids joined together by peptide bonds. An "amino acid" is a subunit that is polymerized to form proteins and there are twenty amino acids that are universally found in proteins. The general formula for an amino acid is $H_2N$-CHR-COOH, in which the R group can be anything from a hydrogen atom (as in the amino acid glycine) to a complex ring (as in the amino acid tryptophan).

A functional portion of an individual component of the complexes may be defined here as a protein

portion of an individual component of a complex still capable of forming a stable complex with another member of the complex under standard cellular and physiological conditions. For example, a functional portion of a component may include, but is not limited to, a protein portion of CCK-4 which is still capable of stably binding a corresponding binding partner domain of an associated protein, and thus is still capable of forming a complex with that protein. Further, in the case of the catalytic domains of the individual components of the invention, a functional portion of a catalytic domain may refer to a protein still capable of stably binding a substrate molecule under standard physiological conditions.

One method utilizing this approach that may be pursued in the isolation of such complex component-binding molecules would include the attachment of a component molecule, or a functional portion thereof, to a solid matrix, such as agarose or plastic beads, micro-titer wells, petri dishes, or membranes composed of, for example, nylon or nitrocellulose, and the subsequent incubation of the attached component molecule in the presence of a potential component-binding compound or compounds. Attachment to said solid support may be direct or by means of a component specific antibody bound directly to the solid support. After incubation, unbound compounds are washed away, component-bound compounds are recovered. By utilizing this procedure, large numbers of types of molecules may be simultane-ously screened for complex component-binding activity.

The complex components which may be utilized in the above screening method may include, but are not limited to, molecules or functional portions thereof, such as catalytic domains, phosphorylation domains, extracellular domains, or portions of extracellular domains, such as ligand-binding domains, and adaptor proteins, or functional portions thereof. The peptides used may be phosphorylated, e.g., may contain at least one phosphorylated amino acid residue, preferably a phosphorylated Tyr amino acid residue, or may be unphosphorylated. A phosphorylation domain may be defined as a peptide region that is specifically phosphorylated at certain amino acid residues. A functional portion of such a phosphorylation domain may be defined as a peptide capable of being specifically phosphorylated at certain amino acids by a specific protein.

Molecules exhibiting binding activity may be further screened for an ability to disrupt protein complexes. Alternatively, molecules may be directly screened for an ability to promote the complexes. For example, in vitro complex formation may be assayed by, first, immobilizing one component, or a functional portion thereof, of the complex of interest to a solid support. Second, the immobilized complex component may be exposed to a compound such as one identified as above, and to the second component, or a functional portion thereof, of the complex of interest. Third, it may be determined whether or not the second component is still capable of forming a complex with the immobilized

component in the presence of the compound. In addition, one could look for an increase in binding.

Additionally, complex formation in a whole cell may be assayed by utilizing co-immunoprecipitation techniques well known to those of skill in the art. Briefly, a cell line capable of forming a complex of interest may be exposed to a compound such as one identified as above, and a cell lysate may be prepared from this exposed cell line. An antibody raised against one of the components of the complex of interest may be added to the cell lysate, and subjected to standard immunoprecipitation techniques. In cases where a complex is still formed, the immunoprecipitation will precipitate the complex, whereas in cases where the complex has been disrupted, only the complex component to which the antibody is raised will be precipitated.

A preferred method for assessing modulation of complex formation within a cell utilizes a method similar to that described above. Briefly, a cell line capable of forming a complex of interest is exposed to a test compound. The cells are lysed and the lysate contacted with an antibody specific to one component of the complex, said antibody having been previously bound to a solid support. Unbound material is washed away, and the bound material is exposed to a second antibody, said second antibody binding specifically to a second component of the complex. The amount of second antibody bound is easily detected by techniques well known in the art. Cells exposed to an inhibitory test compound will have formed a lesser amount of complex compared to cells

not exposed to the test compound, as measured by the amount of second antibody bound. Cells exposed to a test compound that promotes complex formation will have an increased amount of second antibody bound.

The effect of an agent on the differentiation capability of the complex of interest may be directly assayed. Such agents may, but are not required to, include those agents identified by utilizing the above screening technique. For example, an agent or agents may be administered to a cell such as a neuronal cell, capable of forming a complex, for example, which, in the absence of any agent, would not lead to the cell's differentiation. The differentiation state of the cell may then be measured either in vitro or in vivo. One method of measurement may involve observing the amount of neurile growth present.

Agents capable of disrupting complex formation and capable of reducing or inhibiting disorders, which involve the formation of such complexes, or which involve the lack of formation of such complexes, may be used in the treatment of patients exhibiting or at risk for such disorders. A sufficient amount of agent or agents such as those described above may be administered to a patient so that the symptoms of the disease or condition are reduced or eliminated.

**XX. Purification and Production of Complexes**

Described in this Section are methods for the synthesis or recombinant expression of components, or fragments thereof, of the protein complexes of the invention. Also described herein are methods by which cells exhibiting the protein complexes of the invention may be engineered.

**XXI. Purification Methods**

The complexes of the invention may be substantially purified, i.e., may be purified away from at least 90% (on a weight basis), and from at least 99%, if desired, of other proteins, glycoproteins, and other macromolecules with which it is associated. Such purification can be achieved by utilizing a variety of procedures well known to those of skill in the art, such as subjecting cells, tissue or fluid containing the complex to a combination of standard methods, for example, ammonium sulfate precipitation, molecular sieve chromatography, and/or ion exchange chromatography.

Alternatively, or additionally, a complex may be purified by immunoaffinity chromatography using an immunoadsorbent column to which an antibody is immobilized which is capable of binding to one or more components of the complex. Such an antibody may be monoclonal or polyclonal in origin. Other useful types of affinity purification for the protein complex may utilize, for example, a solid-phase substrate which binds the catalytic kinase domain of a protein, or an immobil-

ized binding site for noncatalytic domains of the components of the complex, which bind in such a manner as to not disrupt the complex. The complex of the present invention may be biochemically purified from a variety of cell or tissue sources.

**XXII. Synthesis and Expression Methods**

Methods for the synthesis of polypeptides or fragments thereof, which are capable of acting as components of the complexes of the present invention, are well-known to those of ordinary skill in the art. See, for example, Creighton, Proteins: Structures and Molecular Principles, W.H. Freeman and Co., NY (1983), which is incorporated herein, by reference, in its entirety.

Components of a complex which have been separately synthesized or recombinantly produced, may be reconstituted to form a complex by standard biochemical techniques well known to those skilled in the art. For example, samples containing the components of the complex may be combined in a solution buffered with greater than about 150mM NaCl, at a physiological pH in the range of 7, at room temperature. For example, a buffer comprising 20mM Tris-HCl, pH 7.4, 137mM NaCl, 10% glycerol, 1% Triton X-100, 0.1% SDS, 0.5% deoxycholate and 2mM EDTA could be used.

Methods for preparing the components of complexes of the invention by expressing nucleic acid encoding proteins are described herein. Methods which are well known to those skilled in the art can be used

to construct expression vectors containing protein coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques and in vivo recombination/genetic recombination. DNA and RNA synthesis may, additionally, be performed using an automated synthesizers. See, for example, the techniques described in Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y. (1989), and in Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, N.Y. (1989).

A variety of host-expression vector systems may be utilized to express the coding sequences of the components of the complexes of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the protein complexes of the invention. These include but are not limited to microorganisms such as bacteria (e.g., E.coli, B. subtilis) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing protein coding sequences; yeast (e.g., Saccharomyces and Pichia) transformed with recombinant yeast expression vectors containing the protein coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the protein coding sequences; plant cell systems infected with recombinant virus expression

vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing the protein coding sequences coding sequence; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems a number of expression vectors may be advantageously selected depending upon the use intended for the complex being expressed. For example, when large quantities of complex proteins are to be produced for the generation of antibodies or to screen peptide libraries, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include but are not limited to the E. coli expression vector pUR278 (Ruther et al., *EMBO J.* 2:1791, 1983), in which the protein coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye and Inouye, *Nucleic acids Res.* 13:3101-3109, 1985; Van Heeke & Schuster, *J. Biol. Chem.* 264:5503-5509, 1989); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to

glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned protein can be released from the GST moiety.

In an insect system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The complex coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the PTK/adaptor complex coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect Spodoptera frugiperda cells in which the inserted gene is expressed (e.g., see Smith et al., *J. Biol.* 46:584, 1983; Smith, U.S. Patent No. 4,215,051).

In mammalian host cells, a number of viral based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the complex coding sequence may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by <u>in vitro</u> or <u>in vivo</u> recombination. Insertion in a non-essential region of the viral genome (e.g.,

region E1 or E3) will result in a recombinant virus that is viable and capable of expressing proteins in infected hosts. (E.g., See Logan & Shenk, *Proc. Natl. Acad. Sci. USA* 81:3655-3659, 1984) Specific initiation signals may also be required for efficient translation of inserted coding sequences. These signals include the ATG initiation codon and adjacent sequences.

In cases where an entire protein gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the coding sequence is inserted, exogenous translational control signals, including the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., *Methods in Enzymol.* 153:516-544, 1987)

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the

protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cells lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably coexpress both the proteins may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the protein encoding DNA independently or coordinately controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker.

Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which coexpress

both the PTK and adaptor protein. Such engineered cell lines are particularly useful in screening and evaluation of compounds that affect signals mediated by the complexes.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler et al., *Cell* 11:223, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, *Proc. Natl. Acad. Sci. USA* 48:2026, 1962), and adenine phosphoribosyltransferase (Lowy et al., *Cell* 22:817, 1980) genes can be employed in $tk^-$, $hgprt^-$ or $aprt^-$ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler et al., *Natl. Acad. Sci. USA* 77:3567, 1980; O'Hare et al., *Proc. Natl. Acad. Sci. USA* 78:1527, 1981); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, *Proc. Natl. Acad. Sci.* USA 78:2072, 1981); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin et al., *J. Mol. Biol.* 150:1, 1981); and hygro, which confers resistance to hygromycin (Santerre et al. *Gene* 30:147, 1984) genes.

New members of the protein families capable of forming the complexes of the invention may be identified and isolated by molecular biological techniques well known in the art. For example, a previously unknown protein encoding gene may be isolated by performing a polymerase chain reaction (PCR) using two degenerate oligonucleotide primer pools designed on the basis of

highly conserved sequences within domains common to members of the protein family.

The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from cell lines or tissue known to express complexes. The PCR product may be subcloned and sequenced to insure that the amplified sequences represent the sequences of a member of the PTK or adaptor subfamily. The PCR fragment may then be used to isolate a full length protein cDNA clone by radioactively labeling the amplified fragment and screening a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library. For a review of cloning strategies which may be used. See e.g., Maniatis, Molecular Cloning: A Laboratory Manual, Cold Springs Harbor Press, N.Y. (1989); and Ausubel et al., Current Protocols in Molecular Biology, Green Publishing Associates and Wiley Interscience, N.Y. (1989). A general method for cloning previously unknown proteins has been described by Skolnik (Skolnik, E.Y., *Cell* 65:75, 1991) and Skolnik et al., (U.S. Patent Application Serial No. 07/643,237) which are incorporated herein by reference, in their entirety, including drawings.

**XXIII. Derivatives of Complexes**

Also provided herein are functional derivatives of a complex. By "functional derivative" is meant a "chemical derivative," "fragment," "variant," "chimera," or "hybrid" of the complex, which terms are defined below. A functional derivative retains at least

a portion of the function of the protein, for example reactivity with an antibody specific for the complex, enzymatic activity or binding activity mediated through noncatalytic domains, which permits its utility in accordance with the present invention.

A "chemical derivative" of the complex contains additional chemical moieties not normally a part of the protein. Covalent modifications of the protein complex or peptides are included within the scope of this invention. Such modifications may be introduced into the molecule by reacting targeted amino acid residues of the peptide with an organic derivatizing agent that is capable of reacting with selected side chains or terminal residues, as described below.

Cysteinyl residues most commonly are reacted with alpha-haloacetates (and corresponding amines), such as chloroacetic acid or chloroacetamide, to give carboxymethyl or carboxyamidomethyl derivatives. Cysteinyl residues also are derivatized by reaction with bromotrifluoroacetone, chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide, methyl 2-pyridyl disulfide, p-chloromercuribenzoate, 2-chloromercuri-4-nitrophenol, or chloro-7-nitrobenzo-2-oxa-1,3-diazole.

Histidyl residues are derivatized by reaction with diethylprocarbonate at pH 5.5-7.0 because this agent is relatively specific for the histidyl side chain. Para-bromophenacyl bromide also is useful; the reaction is preferably performed in 0.1 M sodium cacodylate at pH 6.0.

Lysinyl and amino terminal residues are reacted with succinic or other carboxylic acid anhydrides. Derivatization with these agents has the effect or reversing the charge of the lysinyl residues. Other suitable reagents for derivatizing primary amine containing residues include imidoesters such as methyl picolinimidate; pyridoxal phosphate; pyridoxal; chloroborohydride; trinitrobenzenesulfonic acid; O-methylisourea; 2,4 pentanedione; and transaminase-catalyzed reaction with glyoxylate.

Arginyl residues are modified by reaction with one or several conventional reagents, among them phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, and ninhydrin. Derivatization of arginine residues requires that the reaction be performed in alkaline conditions because of the high $pK_a$ of the guanidine functional group. Furthermore, these reagents may react with the groups of lysine as well as the arginine alpha-amino group.

Tyrosyl residues are well-known targets of modification for introduction of spectral labels by reaction with aromatic diazonium compounds or tetranitromethane. Most commonly, N-acetylimidizol and tetranitromethane are used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Carboxyl side groups (aspartyl or glutamyl) are selectively modified by reaction carbodiimide (R'-N-C-N-R') such as 1-cyclohexyl-3-(2-morpholinyl(4-ethyl) carbodiimide or 1-ethyl-3-(4-azonia-4,4-dimethylpentyl) carbodiimide. Furthermore, aspartyl and glutamyl

residue are converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

Glutaminyl and asparaginyl residues are frequently deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Derivatization with bifunctional agents is useful, for example, for cross-linking the component peptides of the complexes to each other or the complex to a water-insoluble support matrix or to other macromolecular carriers. Commonly used cross-linking agents include, for example, 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), and bifunctional maleimides such as bis-N-maleimido-1,8-octane. Derivatizing agents such as methyl-3-[p-azidophenyl) dithiolpropioimidate yield photoactivatable intermediates that are capable of forming crosslinks in the presence of light. Alternatively, reactive water-insoluble matrices such as cyanogen bromide-activated carbohydrates and the reactive substrates described in U.S. Patent Nos. 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; and 4,330,440 are employed for protein immobilization.

Other modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups

of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (Creighton, T.E., Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)), acetylation of the N-terminal amine, and, in some instances, amidation of the C-terminal carboxyl groups.

Such derivatized moieties may improve the stability, solubility, absorption, biological half life, and the like. The moieties may alternatively eliminate or attenuate any undesirable side effect of the protein complex and the like. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co., Easton, PA (1990).

The term "fragment" is used to indicate a polypeptide derived from the amino acid sequence of the proteins, of the complexes having a length less than the full-length polypeptide from which it has been derived. Such a fragment may, for example, be produced by proteolytic cleavage of the full-length protein. Preferably, the fragment is obtained recombinantly by appropriately modifying the DNA sequence encoding the proteins to delete one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence. Fragments of a protein, when present in a complex resembling the naturally occurring complex, are useful for screening for compounds that act to modulate signal transduction, as described below. It is understood that such fragments, when present in a complex may

retain one or more characterizing portions of the native complex. Examples of such retained characteristics include: catalytic activity; substrate specificity; interaction with other molecules in the intact cell; regulatory functions; or binding with an antibody specific for the native complex, or an epitope thereof.

Another functional derivative intended to be within the scope of the present invention is a complex comprising at least one "variant" polypeptide which either lack one or more amino acids or contain additional or substituted amino acids relative to the native polypeptide. The variant may be derived from a naturally occurring complex component by appropriately modifying the protein DNA coding sequence to add, remove, and/or to modify codons for one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence. It is understood that such variants having added, substituted and/or additional amino acids retain one or more characterizing portions of the native complex, as described above.

A functional derivative of complexes comprising proteins with deleted, inserted and/or substituted amino acid residues may be prepared using standard techniques well-known to those of ordinary skill in the art. For example, the modified components of the functional derivatives may be produced using site-directed mutagenesis techniques (as exemplified by Adelman et al., 1983, DNA 2:183) wherein nucleotides in the DNA coding the sequence are modified such that a modified coding sequence is modified, and thereafter expressing

this recombinant DNA in a prokaryotic or eukaryotic host cell, using techniques such as those described above. Alternatively, components of functional derivatives of complexes with amino acid deletions, insertions and/or substitutions may be conveniently prepared by direct chemical synthesis, using methods well-known in the art. The functional derivatives of the complexes typically exhibit the same qualitative biological activity as the native complexes.

Other functional derivatives include mutant, species and allelic variations.

By "mutant variation" is meant a nucleic acid or amino acid molecule that results from any detectable change in the genetic material which may be transmitted to daughter cells giving rise to mutant cells, including nucleic acids or polypeptides having nucleotides or amino acids that are added, deleted, substituted for, inverted, or transposed to new positions with and without inversion. The mutant variation may occur spontaneously or may be induced experimentally by application of mutagens and may result from any (or a combination of) detectable, unnatural change affecting the chemical or physical constitution, mutability, replication, phenotypic function, or recombination of one or more deoxyribonucleotides.

By "species variation" is meant a change in the nucleic acid or amino acid sequence that occurs among species and may be determined by DNA sequencing of the molecule in question.

By "allelic variation" is meant an alternative functional derivation of the typical form of a gene in an organism occupying a given locus on a chromosome.

**XXIV. Evaluation of Disorders**

The protein complexes of the invention involved in disorders may be utilized in developing a prognostic evaluation of the condition of a patient suspected of exhibiting such a disorder. For example, biological samples obtained from patients suspected of exhibiting a disorder involving a protein complex may be assayed for the presence of such complexes. If such a protein complex is normally present, and the development of the disorder is caused by an abnormal quantity of the complex, the assay should compare complex levels in the biological sample to the range expected in normal tissue of the same cell type.

Among the assays which may be undertaken may include, but are not limited to isolation of the protein complex of interest from the biological sample, or assaying for the presence of the complex by exposing the sample to an antibody specific for the complex, but non-reactive to any single, non-complexed component, and detecting whether antibody has specifically bound.

Alternatively, one or more of the components of the protein complex may be present in an abnormal level or in a modified form, relative to the level or form expected is normal, nononcogenic tissue of the same cell type. It is possible that overexpression of both components may indicate a particularly aggressive

disorder. Thus, an assessment of the individual and levels of mRNA and protein in diseased tissue cells may provide valuable clues as to the course of action to be undertaken in treatment of such a disorder. Assays of this type are well known to those of skill in the art, and may include, but are not limited to, Northern blot analysis, RNAse protection assays, and PCR for determining mRNA levels. Assays determining protein levels are also well known to those of skill in the art, and may include, but are not limited to, Western blot analysis, immunoprecipitation, and ELISA analysis. Each of these techniques may also reveal potential differences in the form (e.g., the primary, secondary, or tertiary amino acid sequence, and/or post-translational modifications of the sequence) of the component(s).

**EXAMPLES**

The examples below are non-limiting and are merely representative of various aspects and features of the present invention. The examples below demonstrate the isolation, characterization and expression of CCK-4.

The following cell lines were obtained from the American Type Culture Collection (ATCC): colon carcinoma cell lines SW 480, SW 1463, SW 1417, SW 837, SW 948, SW 620, SW 403, SW 116, T 84, HTC 15, LS 123, HT 29, and Caco-2.

Example 1: PCR amplification and isolation of CCK-4 cDNAs

Poly A+ RNA was prepared from a pool of colon carcinoma cell lines SW 480, SW 1463, SW 1417, SW 837, SW 948, SW 620, SW 403, SW 116, T 84, HTC 15, LS 123, HT 29, and Caco-2 and used as a source to clone cDNA fragments using PCR primers corresponding to the conserved motifs HRDLAA and D(V/M)WS(F/Y)G (Hanks et al., 1988, Wilks, 1989). Approximately 200 PCR clones were analyzed by DNA sequencing, and one clone (designated CCK-4, for colon carcinoma clone 4) was used to screen a custom made l ZAP cDNA library (Clontech) from the same mRNA pool. A 2.2 kb positive clone, designated CCK-4 clone #1, was confirmed by partial DNA sequencing and subsequently used to screen for 5' clones using the same library, as well as additional cDNA libraries from normal tissues including brain (a lgt 11 human caudate nucleus cDNA library - Clontech) and normal human placenta (lZAP human placenta cDNA library - Clontech). cDNA clones from the colon carcinoma pool (2), normal human brain (4), and placenta (4) libraries were analyzed.

Example 2: Characterization of cDNA clones

cDNA clones were purified and characterized using standard molecular procedures (Sambrook et al. 1989). Plasmid DNAs were prepared for sequencing using the Qiagen™ ion exchange resin, according to the Qiagen instructional manual. cDNA clones were sequenced on both strands by double-stranded sequencing using the Taq DNA polymerase cycle and chain terminators (Promega) and fluoroscein labeled primers (Zimmerman et al. 1988), and

analyzed with an Automated Laser Fluorescent Sequencer (Pharmacia LKB). The DNA sequence was also confirmed by conventional sequencing using $^{35}S$ nucleotides and the chain termination method (Sanger et al. 1977). Sequence text files were transferred into the DNA Star () assembly package and analyzed using a variety of programs, including DNA Star and the MacVector sequence analysis package.

Example 3: RNA extraction and Northern analysis.

RNA was extracted from cell lines and poly A$^+$ RNA selected on oligo-dT columns using standard procedures (Sambrook et al. 1989). 1.5 -2 mg of polyA$^+$ RNA was fractionated on 1.2% formaldehyde-containing agarose gels and transferred to nitrocellulose or nytran filters (Schleicher and Schuell). Hybridizations were performed overnight in 50% formamide, 5x SSC (750 mM sodium chloride, 75 mM sodium citrate), 5x Denhardt's (0.1% Ficoll, 0.1% polyvinlypyrollidone, 0.1% BSA) and 50 mM $NaPO_4$ (pH 6.8) at 420C with 1-3x 106 cpm/ml $^{32}P$-random primed DNA cytoplasmic domain probes, followed by high stringency washes in 0.2x SSC, 0.2% SDS at 55$^0$C. The filters were exposed for 11 days.

Northern blot analysis of the human CCK-4 gene was performed. PolyA+ RNAs were prepared from various normal human tissues and colon carcinoma cell lines. Approximately 2 mg of these RNAs were fractionated on formaldehyde-containing agarose gels, transferred to nitrocellulose filters, and hybridized with a probe representing the cytoplasmic domain of CCK-4.

Example 4: Expression constructs

To determine whether the CCK-4 cytoplasmic domain possessed kinase activity chimeric constructs were made which contained the epidermal growth factor receptor (EGF-R)'s extracellular domain fused to the CCK-4's transmembrane and cytoplasmic domains. A 5' oligonucleotide primer (sense strand) (5'-GGATCAGTG**TCTAGA**GCCAACGCCACAACCACCGCGC-3' SEQ. ID. NO. 2) containing an Xba 1 site and mapping to the 5' untranslated region of EGF-R was used in conjunction with a 3' oligonucleotide primer mapping upstream of the transmembrane region (anti-sense strand) (5'-CCCAGTGGC**TTCGAA**CGGGATCTTAGGCCCATTCGTTGG-3' SEQ. ID. NO. 3) and containing a unique BstB 1 site, to amplify the EGF-R extracellular domain. At the same time, the CCK-4 primers includes the 5' (sense strand) BstB 1-containing transmembrane domain primer, (5'-GGATG**TTCGAA**ACCATGGGGTTGTCGGTGGGTGCCGC-3' SEQ. ID. NO. 4) and a hemaglutinin epitope-tagged (Niman et al. 1983; Pati, 1992) (HA1) Not 1-containing 3' oligonucleotide primer (anti-sense primer) mapping at the carboxy-terminal end of the CCK-4 cytoplasmic domain (5'-GGGG**CGGCCGC**TTAAGCGTAATCTGGAACATCGTATGGGTACGGCTTGCTGTCCAC GGTGCTG T-3' SEQ. ID. NO. 5). The primer pairs were used, in conjunction with the relevant templates to amplify the entire CCK-$4_{TM+TK}$ and EGF-REC domains. Following digestion of the EGF-$R_{EC}$ and CCK-$4_{TM+TK}$ PCR products with Xba1/BstB1 and BstB1/Not1, respectively, these fragments were cloned into the cytomegalovirus promoter-based expression vector pCMV [REF] vectors,

and confirmed by restriction analysis and DNA sequencing. HA-tagged full-length CCK-4 clones were similarly constructed.

Example 5: Transient expression of CCK-4 and EC-R

The 293 cell system was used for transient expression of CCK-4. Cells were grown in DMEM supplemented with 10% FCS at 5% $CO_2$ and 37°C. One day prior to transfection $2x10^5$ cells were seeded into each well of a six-well dish. Transfections were carried out according to the protocol of Chen and Okayama (1987) with a total of 4 μg DNA/well. 16 h after transfection, cells were washed once with DMEM and fresh growth medium was added.

For metabolic labeling, cells were grown overnight with [$^{35}$S]methionine (50 μCi/ml) in methionine-free DMEM (0.5 ml/well) containing 1% dialyzed FCS. Lysis of the cells was performed as previously described (Jallal et al., 1992). For immunoprecipitations, either anti-HA antibody or anti-EGF-R antibodies and 20 μl of protein A-sepharose were added to the cleared lysates and incubated at 4°C for 2 hours. Precipitates were washed extensively with HNTG buffer and SDS sample buffer was added to the samples.

Expression of CCK-4, EC-R and EGF-R in transiently transfected 293 cells were studied. 293 cells were transiently transfected with expression plasmids for CCK-4, EC-R, and EGF-R, or the insertless expression vector as a control. After metabolic labeling with [$^{35}$S] methionine cells were lysed and the cleared

lysates subjected to immunoprecipitation with either a mAb directed against the human EGF-R extracellular domain or with a anti-HA antibody. Precipitated proteins were separated on 7.5 % SDS-PAGE and detected by autoradiography of the dried gel.

Example 6: In Vitro Kinase Assay

For the *in vitro* kinase assay, immunoprecipitates were suspended in 10 µl HNTG buffer containing 10 mN manganese chloride. Phosphorylation was carried out at 4°C for 10 min in the presence of 10 µCi [g$^{32}$P]ATP or 1 mM unlabeled ATP. The reaction was stopped by washing the samples with HNTG and the addition of SDS-sample buffer.

Precipitated proteins were separated by SDS-PAGE and either transferred onto nitrocellulose or the gel was processed for direct autoradiographic analysis. Immunoblotting procedures were essentially done as described previously (Jallal et al., 1992).

Analysis of CCK-4 in vitro kinase activity was performed. Lysates from 293 cells transfected with control plasmid or expression plasmids for CCK-4, EC-R, or EGF-R were subjected to immunoprecipitation with antibodies. After extensive washing precipitates were used for *in vitro* kinase reactions. Some precipitates were incubated with [g-32P]ATP. Following SDS-PAGE, phosphorylated proteins were detected by autoradiography. Other precipitates were incubated with 10 mM ATP. Following SDS-PAGE, tyrosine-phosphorylated proteins were detected by immunoblotting with mAb 5E2.

The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

# SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(ii) TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF CCK-4 SIGNAL TRANSDUCTION

(iii) NUMBER OF SEQUENCES:

(iv) CORRESPONDENCE ADDRESS:

| | |
|---|---|
| (A) ADDRESSEE: | Lyon & Lyon |
| (B) STREET: | 633 West Fifth Street |
| (C) CITY: | Los Angeles |
| (D) STATE: | California |
| (E) COUNTRY: | USA |
| (F) ZIP: | 90071 |

(v) COMPUTER READABLE FORM:

| | |
|---|---|
| (A) MEDIUM TYPE: | 3.5" Diskette, 1.44 Mb storage |
| (B) COMPUTER: | IBM Compatible |
| (C) OPERATING SYSTEM: | IBM P.C. DOS (Version 5.0) |
| (D) SOFTWARE: | WordPerfect (Version 5.1) |

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

| | |
|---|---|
| (A) NAME: | Warburg, Richard J. |
| (B) REGISTRATION NUMBER: | 32,327 |
| (C) REFERENCE/DOCKET NUMBER: | 210/248 |

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (213) 489-1600
(B) TELEFAX: (213) 955-0440
(C) TELEX: 67-3510

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(E) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Asp Xaa Gly Xaa Tyr Xaa Cys

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGATCAGTGT CTAGAGCCAA CGCCACAACC ACCGCGC 37

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CCCAGTGGCT TCGAACGGGA TCTTAGGCCC ATTCGTTGG 39

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGATGTTCGA AACCATGGGG TTGTCGGTGG GTGCCGC 37

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGGGCGGCCG CTTAAGCGTA ATCTGGAACA TCGTATGGGT ACGGCTTGCT 50

GTCCACGGTG CTGT 64

What is claimed is:

CLAIMS

1. An isolated, enriched or purified nucleic acid encoding a CCK-4 polypeptide.

2. Thge nucleic acid of claim 1 wherein said nucleic acid is human nucleic acid.

3. The nucleic acid of claim 1 wherein said nucleic acid encodes at least 25 contiguous amino acids of the amino acid sequence shown in Fig. 1.

4. A nucleic acid probe for the detection of a CCK-4 polypeptide in a sample.

5. A recombinant nucleic acid encoding a CCK-4 polypeptide and a vector or a promoter effective to initiate transcription in a host cell.

6. A recombinant nucleic acid comprising a transcriptional region functional in a cell, a sequence complimentary to an RNA sequence encoding a CCK-4 polypeptide and a transcriptional termination region functional in a cell.

7. An isolated, enriched or purified purified CCK-4 polypeptide.

8. An antibody having specific binding affinity to a CCK-4 polypeptide.

9. A hybridoma which produces an antibody having specific binding affinity to a CCK-4 polypeptide.

10. A method of detecting a compound capable of binding to a CCK-4 polypeptide comprising the steps of incubating the compound with a CCK-4 polypeptide and detecting the presence of the compound bound to said CCK-4 polypeptide.

11. Method for treating a patient having a disease or condition characterized by an abnormality in a signal transduction pathway, wherein said signal transduction pathway involves the interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner, comprising the step of disrupting or promoting said interaction in vivo.

12. Method of screening potential agents useful for treatment of a disease or condition characterized by an abnormality in a signal transduction pathway, wherein said signal transduction pathway involves the interaction between a CCK-4 receptor tyrosine kinase and a binding partner for said receptor, comprising the step of assaying said potential agents for those able to disrupt or promote said interaction as an indication of a useful said agent.

13. Method for diagnosis of a disease or condition characterized by an abnormality in a signal transduction pathway, wherein said signal transduction pathway involves the interaction between a CCK-4 receptor tyrosine kinase and a CCK-4 binding partner, comprising the step of detecting the level of said interaction as an indication of said disease or condition.

14. Method for diagnosis of colon cancer comprising the step of detecting the presence of a CCK-4 polypeptide or nucleic acid encoding a CCK-4 polypeptide in a patient suspected of having colon cancer.

15. Method for diagnosis of a subset of colon cancers characterized by the presence of a CCK-4 polypeptide or nucleic acid encoding a CCK-4 polypeptide comprising the step of detecting the presence of a CCK-4 polypeptide or nucleic acid encoding a CCK-4 polypeptide in a patient suspected of having colon cancer.

FIG. 2.

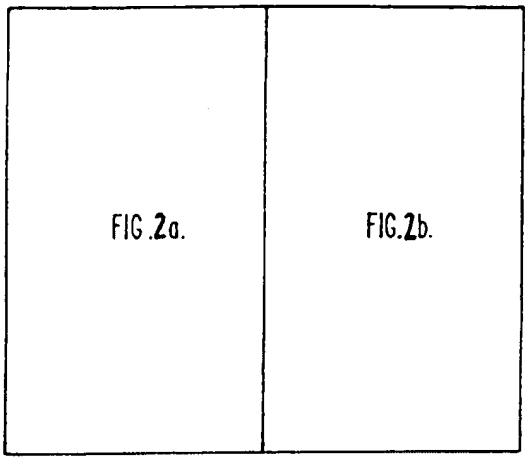


FIG. 1.

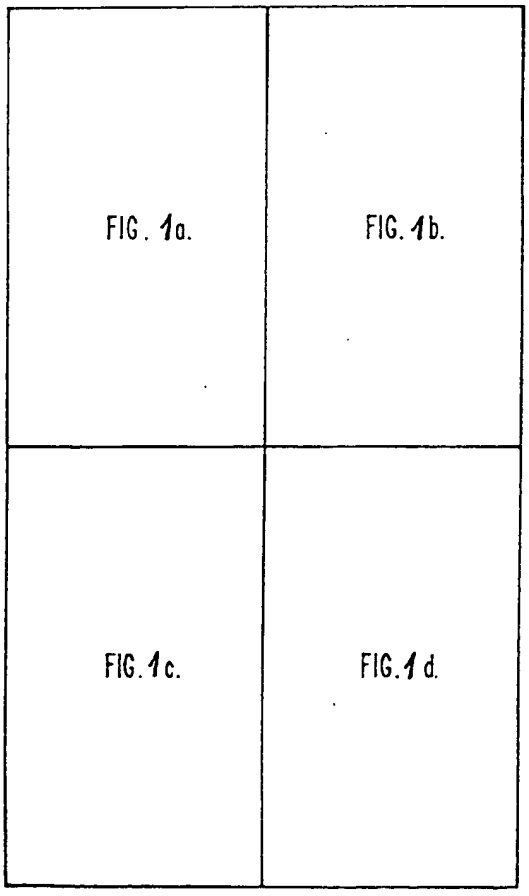

FIG. 1a.

```
   1  CGGGGACTCGGAGGTACTGGGCGCGCGCGG CTCCGGCTCGGGACGCCTCGGGACGCCTCG

   1
 121  CTGTGCCCGCCGCGGAGCAAGTCTGCGCGC CCGCCGTGCGCCCCTAAGCTCCTTTTACCT
      ____________________________
  17  L  S  V  L  L  L  P  L  L  G   G  T  Q  T  A  I  V  F  I  K
 241  CTCAGCGTCCTGCTGCTGCCGCTGCTGGGC GGTACCCAGACAGCCATTGTCTTCATCAAG

  57  A  P  G  P  V  H  V  Y  W  L   L  D  G  A  P  V  Q  D  T  E
 361  GCTCCGGGCCCGGTACATGTGTACTGGCTG CTCGATGGGGCCCCTGTCCAGGACACGGAG
                                                                  +
  97  G  T  F  Q  C  V  A  R  D  D   V  T  G  E  E  A  R  S  A  N
 481  GGCACCTTCCAGTGTGTGGCTCGGGATGAT GTCACTGGAGAAGAAGCCCGCAGTGCCAAC

 137  E  A  E  I  Q  P  Q  T  Q  V   K  L  R  C  H  I  D  G  H  P
 601  GAAGCTGAGATCCAGCCACAGACCCAGGTC AAACTTCGTTGCCACATTGATGGGCACCCT
                      +
 177  T  V  S  S  K  E  R  N  L  I   L  R  P  A  G  P  E  H  S  G
 721  ACAGTCAGCAGCAAGGAGCGGAACCTGACG CTCCGGCCAGCTGGTCCTGAGCATAGTGGG

 217  L  S  I  A  D  E  S  F  A  R   V  V  L  A  P  Q  D  V  V  V
 841  TTGAGCATTGCTGATGAAAGCTTTGCCAGG GTGGTGCTGGCACCCCAGGACGTGGTAGTA
                                  +
 257  Q  W  L  F  E  D  E  T  P  I   T  N  R  S  R  P  P  H  L  R
 961  CAGTGGCTCTTTGAGGATGAGACTCCCATC ACTAACCGCAGTCGCCCCCCACACCTCCGC

 297  G  I  Y  R  C  I  G  Q  G  Q   R  G  P  P  I  I  L  E  A  T
1081  GGGATCTACCGCTGCATTGGCCAGGGGCAG AGGGGCCCACCCATCATCCTGGAAGCCACA

 337  S  E  E  R  V  T  C  L  P  P   K  G  L  P  E  P  S  V  W  W
1201  AGCGAGGAGCGTGTGACCTGCCTTCCCCCC AAGGGTCTGCCAGAGCCCAGCGTGTGGTGG

 377  V  L  A  N  I  A  E  S  D  A   G  V  Y  T  C  H  A  A  N  L
1321  GTGTTGGCCAATATTGCTGAAAGTGATGCT GGTGTCTACACCTGCCACGCGGCCAACCTG

 417  K  P  Q  D  S  Q  L  E  E  G   K  P  G  Y  L  D  C  L  T  Q
1441  AAGCCCCAAGACAGCCAGCTGGAGGAGGGC AAACCCGGCTACTTGGATTGCCTGACCCAG
                      +
 457  R  F  E  V  F  K  N  G  T  L   R  I  N  S  V  E  V  Y  D  G
1561  CGGTTCGAGGTCTTCAAGAATGGGACCTTG CGCATCAACAGCGTGGAGGTGTATGATGGG

 497  Q  V  L  E  K  L  K  F  T  P   P  P  Q  P  Q  Q  C  M  G  F
1681  CAAGTGCTGGAAAAGCTCAAGTTCACACCA CCACCCCAGCCACAGCAGTGCATGGGGTTT

 537  E  R  A  D  G  S  S  L  P  E   W  V  T  D  N  A  G  T  L  H
1801  GAACGGGCAGATGGGAGCAGCCTCCCAGAG TGGGTGACAGACAACGCTGGGACCCTGCAT
```

FIG. 1b.

```
GGGTCGGGTTCCGGTTGCGGCTGCTGCTGC GGCGCCCGCGCTTCCGTAGCGTTCCGCCTC

            M  G  A  A  R  G     S  P  A  R  P  R  R  L  P  L
GAGCCCGCCGCGATGGGAGCTGCGCGGGGA TCCCCGGCCAGACCCCGCCGGTTGCCTCTG

Q  P  S  S  Q  D  A  L  Q  G     R  R  A  L  L  R  C  E  V  E
CAGCCGTCCTCCCAGGATGCACTGCAGGGG CGCCGGGCGCTGCTTCGCTGTGAGGTTGAG

R  R  F  A  Q  G  S  S  L  S     F  A  A  V  D  P  L  Q  D  S
CGGCGTTTCGCCCAGGGCAGCAGCCTGAGC TTTGCAGCTGTGGACCCGCTGCAGGACTCT

A  S  F  N  I  K  W  I  E  A     G  P  V  V  L  K  H  P  A  S
GCCTCCTTCAACATCAAATGGATTGAGGCA GGTCCTGTGGTCCTGAAGCATCCAGCCTCG
                                                          +
R  P  T  Y  Q  W  F  R  D  G     T  P  L  S  D  G  Q  S  N  H
CGGCCCACCTACCAATGGTTCCGAGATGGG ACCCCCCTTTCTGATGGTCAGAGCAACCAC
                                                      +
L  Y  S  C  C  A  H  S  A  F     S  Q  A  C  S  S  Q  N  F  T
CTGTATTCCTGCTGCGCCCACAGTGCTTTT AGCCAGGCTTGCAGCAGCCAGAACTTCACC

A  R  Y  E  E  A  M  F  H  C     Q  F  S  A  Q  P  P  P  S  L
GCGAGGTATGAGGAGGCCATGTTCCATTGC CAGTTCTCAGCCCAGCCACCCCCGAGCCTG
                +
R  A  T  V  F  A  N  G  S  L     L  L  T  Q  V  R  P  R  N  A
AGAGCCACAGTGTTTGCCAACGGGTCTCTG CTGCTGACCCAGGTCCGGCCACGCAATGCA

L  H  L  A  E  I  E  D  M  P     L  F  E  P  R  V  F  T  A  G
CTTCACCTAGCAGAGATTGAAGACATGCCG CTATTTGAGCCACGGGTGTTTACAGCTGGC

E  H  A  G  V  R  L  P  T  H     G  R  V  Y  Q  K  G  H  E  L
GAGCACGCGGGAGTCCGGCTGCCCACCCAT GGCAGGGTCTACCAGAAGGGCCACGAGCTG
                   +
A  G  Q  R  R  Q  D  V  N  I     T  V  A  T  V  P  S  W  L  K
GCTGGTCAGCGGAGACAGGATGTCAACATC ACTGTGGCCACTGTGCCCTCCTGGCTGAAG

A  T  P  K  P  T  V  V  W  Y     R  N  Q  M  L  I  S  E  D  S
GCCACACCAAAACCTACAGTTGTCTGGTAC AGAAACCAGATGCTCATCTCAGAGGACTCA

T  W  Y  R  C  M  S  S  T  P     A  G  S  I  E  A  Q  A  V  L
ACATGGTACCGTTGTATGAGCAGCACCCCA GCCGGCAGCATCGAGGCGCAAGCCGTGCTC

D  K  E  A  T  V  P  C  S  A     T  G  R  E  K  P  T  I  K  W
GACAAGGAGGCCACGGTGCCCTGTTCAGCC ACAGGCCGAGAGAAGCCCACTATTAAGTGG
                                  +
F  A  R  V  T  R  D  D  A  G     N  Y  T  C  I  A  S  N  G  P
TTTGCCCGGGTGACTCGAGATGACGCTGGC AACTACACTTGCATTGCCTCCAACGGGCCG
```

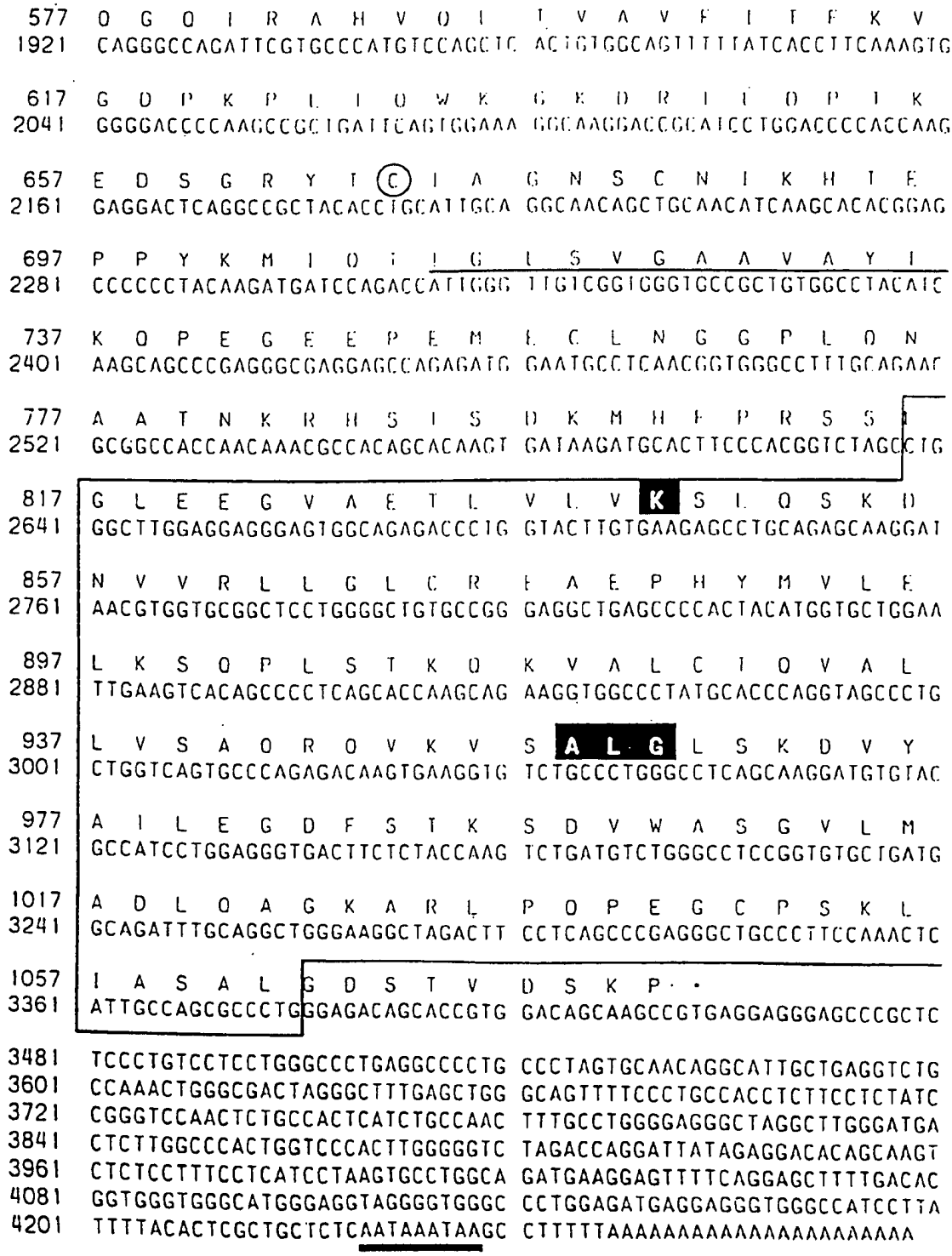
577 Q G Q I R A H V Q L T V A V F I T F K V
1921 CAGGGCCAGATTCGTGCCCATGTCCAGCTC ACTGTGGCAGTTTTTATCACCTTCAAAGTG
617 G D P K P L I Q W K G K D R I I Q P I K
2041 GGGGACCCCAAGCCGCTGATTCAGTGGAAA GGCAAGGACCGCATCCTGGACCCCACCAAG
657 E D S G R Y I C I A G N S C N I K H I E
2161 GAGGACTCAGGCCGCTACACCTGCATTGCA GGCAACAGCTGCAACATCAAGCACACGGAG
697 P P Y K M I Q I I G L S V G A A V A Y I
2281 CCCCCCTACAAGATGATCCAGACCATTGGG TTGTCGGTGGGTGCCGCTGTGGCCTACATC
737 K Q P E G E E P E M E C L N G G P L Q N
2401 AAGCAGCCCGAGGGCGAGGAGCCAGAGATG GAATGCCTCAACGGTGGGCCTTTGCAGAAC
777 A A T N K R H S I S D K M H F P R S S I
2521 GCGGCCACCAACAAACGCCACAGCACAAGT GATAAGATGCACTTCCCACGGTCTAGCCTG
817 G L E E G V A E T L V L V K S I Q S K D
2641 GGCTTGGAGGAGGGAGTGGCAGAGACCCTG GTACTTGTGAAGAGCCTGCAGAGCAAGGAT
857 N V V R L L G L C R E A E P H Y M V L E
2761 AACGTGGTGCGGCTCCTGGGGCTGTGCCGG GAGGCTGAGCCCCACTACATGGTGCTGGAA
897 L K S Q P L S T K Q K V A L C I Q V A L
2881 TTGAAGTCACAGCCCCTCAGCACCAAGCAG AAGGTGGCCCTATGCACCCAGGTAGCCCTG
937 L V S A Q R Q V K V S A L G L S K D V Y
3001 CTGGTCAGTGCCCAGAGACAAGTGAAGGTG TCTGCCCTGGGCCTCAGCAAGGATGTGTAC
977 A I L E G D F S T K S D V W A S G V L M
3121 GCCATCCTGGAGGGTGACTTCTCTACCAAG TCTGATGTCTGGGCCTCCGGTGTGCTGATG
1017 A D L Q A G K A R L P Q P E G C P S K L
3241 GCAGATTTGCAGGCTGGGAAGGCTAGACTT CCTCAGCCCGAGGGCTGCCCTTCCAAACTC
1057 I A S A L G D S T V D S K P ·
3361 ATTGCCAGCGCCCTGGGAGACAGCACCGTG GACAGCAAGCCGTGAGGAGGGAGCCCGCTC
3481 TCCCTGTCCTCCTGGGCCCTGAGGCCCCTG CCCTAGTGCAACAGGCATTGCTGAGGTCTG
3601 CCAAACTGGGCGACTAGGGCTTTGAGCTGG GCAGTTTTCCCTGCCACCTCTTCCTCTATC
3721 CGGGTCCAACTCTGCCACTCATCTGCCAAC TTTGCCTGGGGAGGGCTAGGCTTGGGATGA
3841 CTCTTGGCCCACTGGTCCCACTTGGGGGTC TAGACCAGGATTATAGAGGACACAGCAAGT
3961 CTCTCCTTTCCTCATCCTAAGTGCCTGGCA GATGAAGGAGTTTTCAGGAGCTTTTGACAC
4081 GGTGGGTGGGCATGGGAGGTAGGGGTGGGC CCTGGAGATGAGGAGGGTGGGCCATCCTTA
4201 TTTTACACTCGCTGCTCTCAATAAATAAGC CTTTTTAAAAAAAAAAAAAAAAAAAAAA

FIG. 1c.

E P E R T T V Y Q G H I A L L Q C E A Q
GAACCAGAGCGTACGACTGTGTACCAGGGC CACACAGCCCTACTGCAGTGCGAGGCCCAG

L G P R M H I F Q N G S L V I H D V A P
CTGGGACCCAGGATGCACATCTTCCAGAAT GGCTCCCTGGTGATCCATGACGTGGCCCCT

A P L Y V V D K P V P E E S E G P G S P
GCCCCCCTCTATGTCGTGGACAAGCCTGTG CCGGAGGAGTCGGAGGGCCCTGGCAGCCCT

I A V L G L M F Y C K K R C K A K R L Q
ATTGCCGTGCTGGGCCTCATGTTCTACTGC AAGAAGCGCTGCAAAGCCAAGCGGCTGCAG

G Q P S A E I Q E E V A L T S L G S G P
GGGCAGCCCTCAGCAGAGATCCAAGAAGAA GTGGCCTTGACCAGCTTGGGCTCCGGCCCC

Q P I T T L **G K S E F G** E V F L A K A Q
CAGCCCATCACCACGCTGGGGAAGAGTGAG TTTGGGGAGGTGTTCCTGGCAAAGGCTCAG

E Q Q Q L D F R R E L E M F G K L N H A
GAGCAGCAGCAGCTGGACTTCCGGAGGGAG TTGGAGATGTTTGGGAAGCTGAACCACGCC

Y V D L E D L K Q F L R I S K S K D E K
TATGTGGATCTGGAAGACCTCAAGCAGTTC CTGAGGATTTCCAAGAGCAAGGATGAAAAA

G M E H L S N N R F V **H K D L A A R N** C
GGAATGGAGCACCTGTCCAACAACCGCTTT GTGCATAAGGACTTGGCTGCGCGTAACTGC

N S E Y Y H F R Q A W V A L R W M S P E
AACAGTGAGTACTACCACTTCCGCCAGGCC TGGGTGGCGCTGCGCTGGATGTCCCCCGAG

W E V F T H G E M P H G G Q A D D E V L
TGGGAAGTGTTTACACATGGAGAGATGCCC CATGGTGGGCAGGCAGATGATGAAGTACTG

**Y R L M** Q R C W A L S P K D R P S F S E
TATCGGCTGATGCAGCGCTGCTGGGCCCTC AGCCCCAAGGACCGGCCCTCCTTCAGTGAG

AGGATGGCCTGGGCAGGGGAGGACATCTCT AGAGGGAAGCTCACAGCATGATGGGCAAGA

AGCaGGGCCTGGCCTTTCCTCCTCTTCCTC ACCCTCATCCTTTGGGAGGCTGACTTGGAC
AGGGACAGTGTGGGTGCCACAGGTAACCCC AATTTCTGGCCTTCAACTTCTCCCCTTGAC
GCTGGGTTTGTGGGGAGTTCCTTAATATTC TCAAGTTCTGGGCACACAGGGTTAATGAGT
GAGTCCTCCCCACTCTGGGCTTGTGCACAC TGACCCAGACCCACGGTCTTCCCCACCCTT
TATATAAACCGCCCTTTTTGTATGCACCAC GGGCGGCTTTTATATGTAATTGCAGCGTGG
CCCCACACTTTTATTGTTGTCGTTTTTTGT GTGTTTGTGTTTTTTTGTTTTTGTTTTTGT

FIG. 1d

```
1    MGAARGSPARPRRLPLLSVLLLPLLGGTQT
1    MAA----------LRALLLLLLAVGAQA

61   VHVYWLLDGAPVQDTERRFAQGSSLSFAAV
48   VEFEWLQNGLPIQDTEQRFKEGSNLQFAAV

121  IKWIEAGPVVLKHPASEAEIQPQTQVKLRC
108  IKWMETGSVVLKQPASAAEIQPSSTVVLRC

181  KERNLTLRPAGPEHSGLYSCCAHS-AFSQA
168  KERTLTLRGAGPDDNGLYYCSARPRAVGSV

240  EEAMFHCQFSAQPPPSLQWLFEDETPITNR
228  EEAMFDCQFAAVPPPTQEWLFEDS-PITNR

300  RCIGQGQRGPPIILEATLHLAEIEDMPLFE
281  KCIGHGQKGKALVLKATLRLAEIEEMAPFS

360  GVRLPTHGRVYQKGHELVLANIAESDAGVY
341  QERVPTAGRVYQEAEQLVFTSITEADAGIY

420  DSQLEEGKPGYLDCLTQATPKPTVVWYRNQ
401  DSQLEESKPGYLHCLSKASLKPTVTWYRNG

480  RCMSSTPAGSIEAQAVLQVLEKLKFTPPPQ
461  KCVSSTPAGSIEGYARVHVLEKLKFTPPPQ

540  DGSSLPEWVTDNAGTLHFARVTRDDAGNYT
521  DGSSLPSHVSHRAGILSFHKVSRSDSGNYT

600  RTTVYQGHTALLQCEAQGDPKPLIQWKGKD
581  PTTVYQGHTAMFQCQAEGDPVPHIQWKGKD

660  GRYTCIAGNSCNIKHTEAPLYVVDKPVPEE
641  GKYTCIAGNSCNIKHREAFLYVVDKPAAEE

720  LGLMFYCKKRCKAKRLQKQPEGEEPEMECL
701  LGLMFYCKKRRKAKRLKKHPEGEEPEMECL

779  TNKRHSTSDKMHFPRSSLQPITTLGKSEFG
761  S-KRHSARDKMHFPRSNLQTITTLGRGEFG

839  QQLDFRRELEMFGKLNHANVVRLLGLCREA
820  LQLDFRREAEMFGKLNHVNVVRLLGLCREA

899  SQPLSTKQKVALCTQVALGMEHLSNNRFVH
880  PQPLSTKHKVSLCTQVALGMEHLSNGRFVH

959  EYYHFRQAWVALRWMSPEAILEGDFSTKSD
940  EYYHFRQAWIPLRWMPPEAVLEDEFSTKSD

1019 LQAGKARLPQPEGCPSKLYRLMQRCWALSP
1000 LKSGKTKLPQPEGCPSRLTKLMQRCWAPSP
```

FIG. 2a

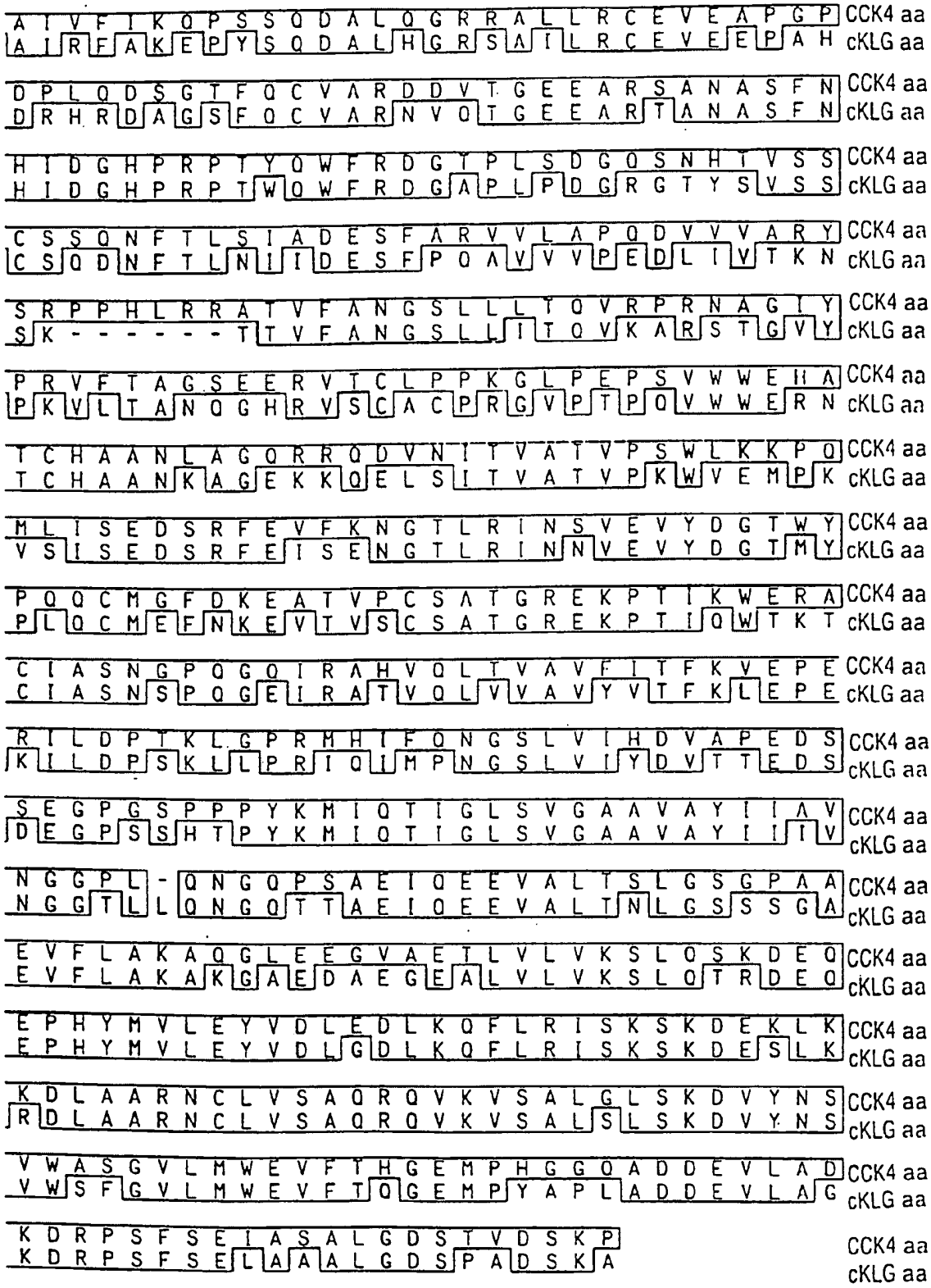


FIG. 2b

SUBSTITUTE SHEET (RULE 26)